medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Stochastic simulation of successive waves of COVID-19 in the province of Barcelona.

M. Bosman[1*], A. Esteve[2,3], L. Gabbanelli[1], X. Jordan[4], A. López-Gay[2,5], M. Manera[1,6], M. Martínez[1,7], P. Masjuan[1,8], Ll.M. Mir[1], J. Paradells[4,9], A. Pignatelli[1], I. Riu[1], V. Vitagliano[1,10,11]

[1] *Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Barcelona, Spain*

[2] *Centre d'Estudis Demogràfics (CED), Barcelona, Spain*

[3] *Serra Húnter Fellow, Departament de Ciències Polítiques i Socials, Universitat Pompeu Fabra, Barcelona, Spain*

[4] *i2CAT Foundation, Edifici Nexus (Campus Nord UPC), Barcelona, Spain*

[5] *Departament de Geografia, Universitat Autònoma de Barcelona, Bellaterra, Spain*

[6] *Serra Húnter Fellow, Departament de Física, Universitat Autònoma de Barcelona, Bellaterra, Spain*

[7] *Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain*

[8] *Departament de Física, Universitat Autònoma de Barcelona, Bellaterra, Spain*

[9] *Departament d'Enginyeria Telemàtica, Universitat Politècnica de Catalunya, Barcelona, Spain*

[10] *Department of Mathematics and Physics, University of Hull, Kingston upon Hull, HU6 7RX, UK*

[11] *DIME, University of Genova, Via all'Opera Pia 15, 16145 Genova, Italy*

**Corresponding author: bosman@ifae.es*

Abstract

Analytic compartmental models are currently used in mathematical epidemiology to forecast the COVID-19 pandemic evolution and explore the impact of mitigation strategies. In general, such

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

models treat the population as a single entity, losing the social, cultural and economical specificities. We present a network model that uses socio-demographic datasets with the highest available granularity to predict the spread of COVID-19 in the province of Barcelona. The model is flexible enough to incorporate the effect of containment policies, such as lockdowns or the use of protective masks, and can be easily adapted to future epidemics. We follow a stochastic approach that combines a compartmental model with detailed individual microdata from the population census, including social determinants and age-dependent strata, and time-dependent mobility information. We show that our model reproduces the dynamical features of the disease across two waves and demonstrate its capability to become a powerful tool for simulating epidemic events.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Introduction

The COVID-19 outbreak struck the society like a tsunami, affecting citizens in many aspects, from mental health to economic well being. Communities all over the world had to modify habits and social contacts, in a common effort of erecting a barricade against the virus. Understanding the virus, its structure and its dynamics, is of paramount importance not only for the development of medical solutions (vaccines and specific pharmacological therapies) but also to define the social strategy of prophylaxis and prevent the expansion of the pathology.

The contagious spread of COVID-19 and the impact of mitigation strategies are currently the subject of many studies, all of them pursuing a better understanding of the variables at stake (it is worth citing the progenitor studies by Imperial College London [1, 2] and, in the particular case of Catalonia, those conducted by the Computational Biology and Complex Systems group at UPC [3, 4, 5]; for further references see also the review by Estrada [6]). This is a complex problem because it involves several factors acting simultaneously, with different weights and consequences, and both short- and long-term effects. What we know for certain is that the disease propagates mainly through social contacts [7]. It becomes then crucial to understand the pattern of contacts of each individual – and how this is intertwined with demographic and social determinants – to comprehend the natural history of COVID-19, as well as the specific attributes (age, previous medical conditions, etc) that determine the outcome of the infection in each infected person [8]. The frequencies and types of contacts are influenced by lockdown policies, while the probability to become infected depends on individual protection measures. Additionally, the probability for

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

infected people to be diagnosed depends not only on the severity of the symptoms but also on testing policies in place. Global epidemiological indicators, like the incidence of the disease, hide a complex entanglement of highly diversified social and individual characteristics that perhaps can be ignored in a first approximation, but ends up being indispensable to model the epidemic.

Several approaches can be pursued. Monitoring indicators, like the daily number of positive tests [3, 4, 5], are useful to track the dynamical evolution of the outbreak in varying conditions and allow simple short-term extrapolations. These results, however, are very sensitive to initial conditions and need a continuous readjustment of inputs to be reliable in the long run. The exponential amplification of small perturbations in the initial data, typical of epidemics, makes the outcome of an outbreak intrinsically unpredictable [9]. As for similar forecasting models, epidemic models cannot return exact predictions; epidemic models can, at most, analyse the likelihoods of different scenarios [10].

In this context, simulation tools turn out to be decisive to explore *in silico* the phase space of all the known variables and to unravel those steps in the contagion whose interdiction would prevent the epidemic progression (for example, by government policies regarding the introduction of tracking software, quarantine of exposed individuals, selective lockdown or mandatory use of protective masks in certain conditions). The purpose of such tools is thus threefold: first, they are *monitoring* tools, able to catch the evolution of the disease; second, they are *forecasting* tools, used to identify the conditions that could favour an outbreak; finally, they are *prevention* tools that can provide health authorities with studies on the effectiveness of containment measures. Analytical

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

models as simulation tools usually treat the population as an aggregated body without taking into account its spatial variability or its social and demographic diversity. Some relevant features of COVID-19 can still be inferred from more sophisticated mathematical models (for Catalonia, for example, this has been done following a microscopic Markov chain approach applied to an age-stratified meta-population model [11]). The loss of granularity, however, precludes the design of fine-tuned policies, which in turn renders the decision-making process not fully adequate in all aspects and for all groups, as proved not long ago for the case of H5N1 influenza [12].

In this work we present a different approach, a simulation tool that follows the daily contact network history of each individual in a largely populated area, the province of Barcelona. A precise knowledge of the morphology of these networks has been shown to be of the greatest importance to capture the salient characteristics of the outbreak [13, 14]. For each individual, we consider their corresponding microenvironment: home location and the co-residence structure, the employment situation and the mobility routine with its resulting pattern of contacts. Using such information, we have developed a stochastic compartmental model that includes both population and epidemiological data. On the one hand, population data consist of all information about contact networks and mixing patterns (the "where" and "how" people live and move), based on the following external inputs:

i Socio-demographic individual microdata from the latest census data provided by the Spanish National Statistics Institute [15].

ii Contacts matrices for densely populated (and country-dependent) environments [16, 17, 18].

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

iii Mobility data, reconstructed from information supplied by mobile network operators [19, 20].

On the other hand, epidemiological data [8] contain information related to the disease itself, its dynamics, and the different phases of an infectious process (exposure to the virus, infection, diagnosis and recovery). In this regard, we use a compartmental model, where the entire population is classified into five states: "susceptible", "exposed", "infected", "diagnosed" and "recovered" (see **Methods**).

Our simulation tool calculates the daily probabilities for each individual of becoming infected at home, at work or school, during further social contacts (generically labelled as "Community") and in public transport. Age, co-residence patterns, work activity, local mobility and the usage of public transportation are used to estimate the contacts network of each individual. Measures modifying the network and the frequency of the contacts (lockdown, teleworking, etc) are also taken into account, as well as the introduction of protective masks that modify viral load transmission (see **Methods**).

Results

Our goal is to simulate the COVID-19 outbreak within the population of the province of Barcelona (5.5 million inhabitants) and reproduce the first two epidemic waves experienced in the region during February-June and July-December 2020 for a total of 300 days. We assume the existence of an initial set of about 50 exposed people (0.001% of the population), randomly distributed across the region. The actual location of people changes throughout the day and throughout the week: the

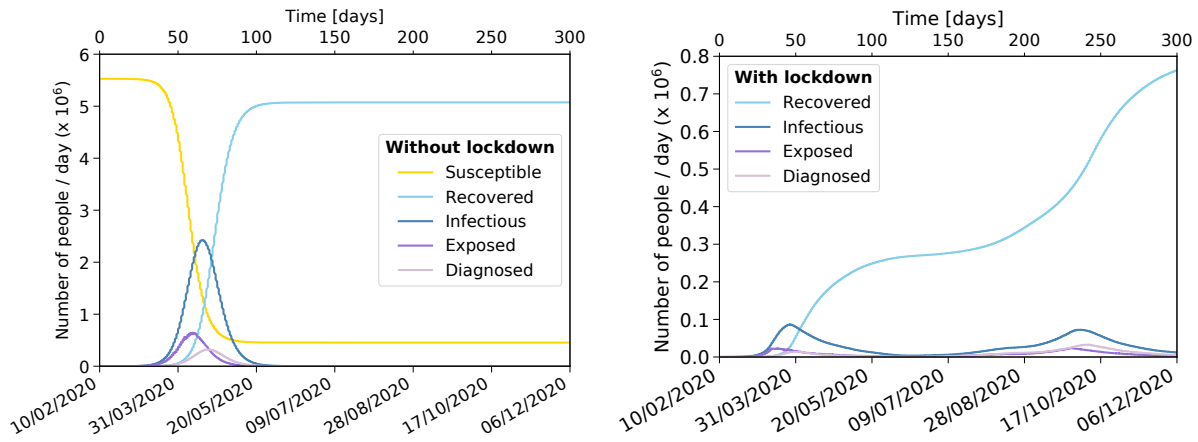
medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

weekdays of workers and students, for example, are divided into three 8-hour fractions according to a given pattern (home $\Rightarrow$ job (or school) $\Rightarrow$ other activities). The pattern of allowed locations and activities evolves according to lockdown and mobility restrictions promulgated during such 300 days.

The number of individuals in each epidemiological compartment predicted by the simulation as a function of time is shown in Figure 1. Without lockdown measures (Figure 1a), more than 90% of the total population of Barcelona and its province get affected by the disease in about 100 days. Taking into account lockdown measures (Figure 1b), the amount of infected people reduces to 300k in about the same period. In the second wave, continuing until the end of the year, the total number of infected people increases to 750k. Figure 2 presents the numbers of currently infected and newly diagnosed people predicted by the simulation as a function of time. The figure presents cumulative and age-dependent evolution curves. The subset of hosts of collective accommodation institutions (nursing homes) is shown separately. The cumulative curve shows a first wave characterised by a steep onset, interrupted by the strict lockdown measures established on March 16, 2020. The mobility restrictions induce the turn-over of the curve and its subsequent decay over a time scale related to the duration of the incubation and infectious phases of the disease. Contacts start increasing again during summer, leading to a second wave in autumn. Adults constitute the most relevant fraction of the infected compartment during both waves. The category of diagnosed is instead dominated by seniors in the first wave and adults in the second.

Relevant results have been obtained by considering separately age-, social-activity- and

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .



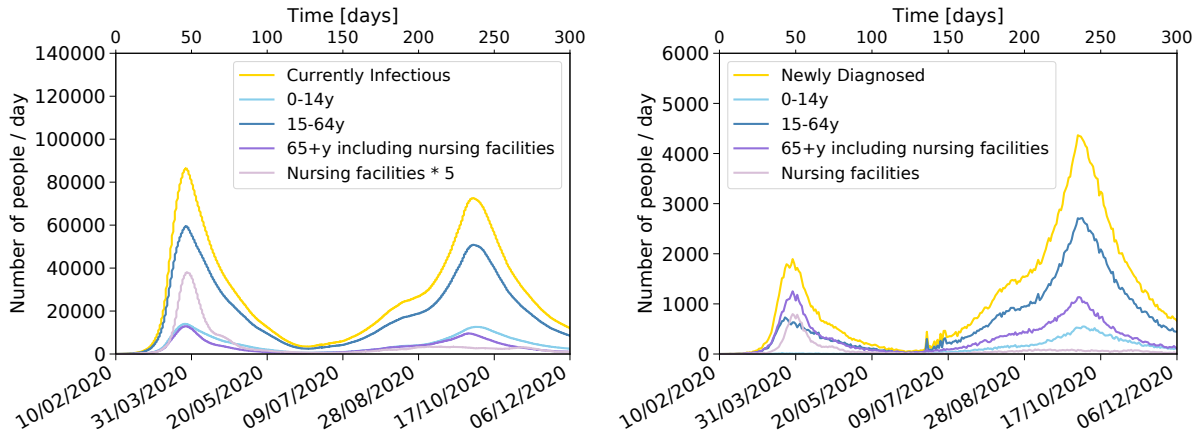
(a) (b)

Figure 1: Evolution of the COVID-19 pandemic in the province of Barcelona. Number of susceptible, exposed, infected, diagnosed and recovered people as a function of time over a period of 300 days from February 10 to December 6, 2020. It has been assumed that 50 people, chosen at random, are exposed at time zero. **a** No lockdown measures. **b** Lockdown measures are applied [susceptible compartment is omitted].

region-dependence:

i The number of diagnosed people, as shown in Figure 2b, can be compared with data provided by the Catalan Health Service [4], as shown in Figure 3. The outcome of the simulation resembles quite closely the real picture emerged in these months for the number of positive PCR tests. Seniors in nursing facilities are notably more numerous than expected from their natural demographic fraction. This pattern results from the joint action of three factors. First, the probability of developing high viral load with strong symptoms increases significantly as a function of age starting at 65 years-old, while adults have a much lower probability to develop strong symptoms, and even less so children. Second, tightly knitted communities as

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .



(a) (b)

Figure 2: Time evolution of infected people. The number of infected people over a period of 300 days is shown split by age categories: less than 15 years old, from 15 to 64 years old and 65 years old and above (hosts in nursing facilities are shown separately). It has been assumed that 50 people, chosen at random, are exposed at time zero. **a** Total number of infected people in a given day. The number of people in nursing facilities is multiplied by five to make it more visible. **b** Newly diagnosed people every day. In this case, the number of people in nursing facilities is not multiplied by five.

those of nursing facilities favour multiple contacts. Third, during the first wave people with strong symptoms would reach the public health system faster and have a higher probability of being diagnosed. The features of the second wave with a peak in October are also qualitatively reproduced, with the spread of the disease picking up again during summer. In this case adults form the group dominating the diagnosed compartment, with children further contributing visibly. The special protection measures taken in nursing facilities (wearing protective masks) reduced contagion, while the amount of tests (and hence the probability to get diagnosed) increased and reached a larger part of the population. Our simulation shows

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

that the unusual strength of the second wave has its root in extra occasional contacts taking place during the summer.

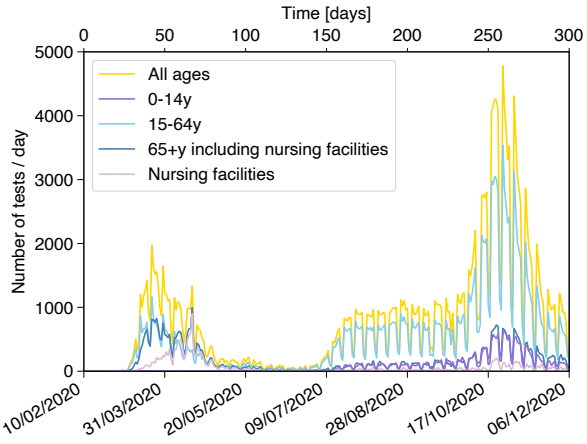


Figure 3: Number of positive PCR tests in the province of Barcelona. The number of positive PCR tests in the province of Barcelona for the period between February 10 and December 6, 2020 is shown separately for the different age segments of the population, including its sum.

ii We analysed under which circumstances people become infected either being at home, at work or school, or involved in social activities, such as shopping, leisure time, etc. Public transportation effects are taken into account: mostly during work time for those using mass transportation for their route to work, or during the weekends or the week for those citizens using group travel systems for other activities (see **Methods**). In our simulation, the amount of available environmental virus, or the viral load each individual is exposed to, is encoded in the *force of infection* λ, an aggregate parameter accounting for the probability to change the status of an individual from susceptible to exposed (see Equation 1 in **Methods**). The results in terms of the viral load are shown in Figure 4. This is an effective way to forecast the risk of contagion in the different social activities. The figure shows that in the early days

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

before lockdown, the viral load exposition is roughly equal at home and at work or school, being slightly higher for community activities. However, when strong lockdown measures are applied on March 16, the contributions from work and community drop drastically and contagion at home becomes the most relevant contribution. The exposition to viral load in public transport or due to casual contacts during summer activities is relatively small, overall of the order of 1%, but is nevertheless a relevant component as it can bring the virus to households and work communities that would otherwise not be exposed to the infection.

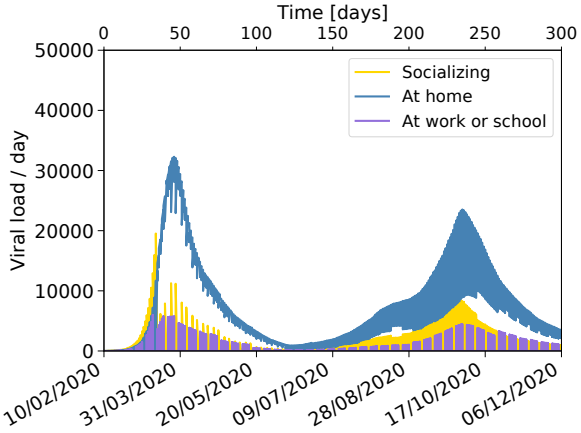
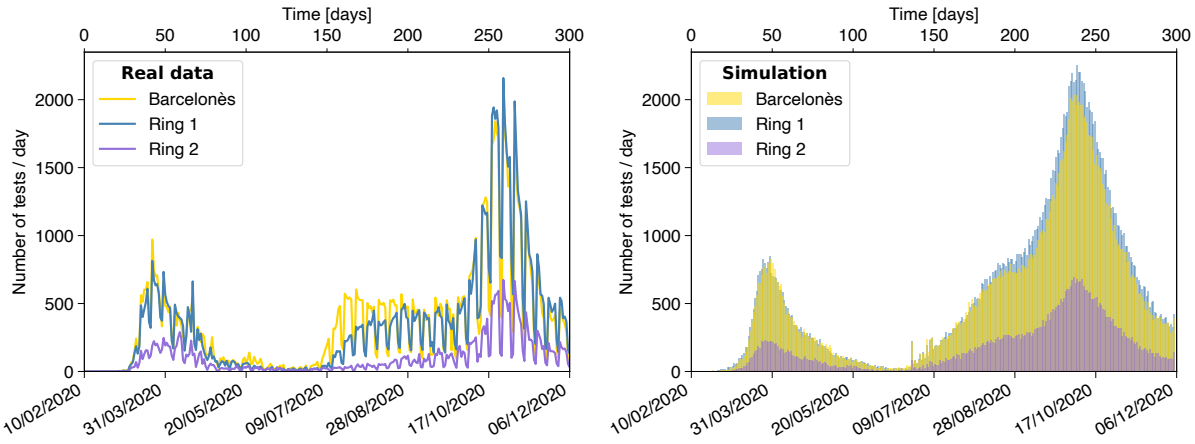


Figure 4: Total viral load evolution. The evolution of the main components of the viral load to which the population is exposed is shown separately for time spent at home, at work or school, or in social/community interaction. It has been assumed that 50 people, chosen at random, are exposed on February 10. The time evolution is shown for 300 days from February 10 to December 6, 2020.

iii We investigate the spatial dependence within our simulation. Real data [4] and the corresponding simulation are shown in Figure 5a and Figure 5b, respectively, for three regions regrouped in three concentric areas (see **Supplementary Material**). At the center, the Barcelonès – 2.3 million inhabitants – exhibits a rather sharp increase of the number of

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .



(a) (b)

Figure 5: Number of positive PCR tests in the province of Barcelona. The results are shown separately for the Barcelonès county, and the counties surrounding it for the first and the second ring. **a** Real data. **b** Simulation.

diagnosed people in early July, and then a plateau until the development of the peak of the second wave. A first "ring" – 2.5 million inhabitants – shows a somewhat slower onset of the second wave with respect to the Barcelonès, with a similar clear peak for the second wave. A second "ring" – 0.8 million inhabitants –, exhibits a similar pattern with an even smaller onset during July and a less intense peak during October. The detailed features of the rise are related to the corresponding pattern of summer contacts. In our simulation we implemented a simple common pattern in all regions. The resulting rise is in average slower than in the data. On the other hand the relative size of the second peak appearing simultaneously in the three regions and related to resuming activities in September is reasonably well reproduced although it develops about two weeks earlier than in data.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Discussion

The first and second waves are of quite different nature [21]. The first one is the result of the initially free propagation of the virus, later controlled by the application of strict lockdown measures. Diagnostics were strongly related to the appearance of symptoms. During the second wave, mobility restrictions were alternatively relaxed and tightened depending on the epidemiological situation. The use of masks proved to be essential, as well as the increase of the efficiency of diagnosis of asymptomatic individuals and the corresponding quarantine measures.

The simulation tool presented here is able to reproduce the main features of the first and second waves of the COVID-19 outbreak in Barcelona and its province. The detailed description of two extremely heterogeneous categories of data, the sociodemographic and epidemiological ones, is a crucial ingredient. The sociodemographic individual microdata, convoluted with mobility information, are the base for modelling contacts and are necessary in order to correctly reproduce the expansion of the disease governed by the different components of the parameter λ (Figure 4). Indeed, since the contribution coming from the viral load components linked to work- and community-activities is drastically reduced as soon as lockdown measures prevail, co-residence contagion becomes the dominant factor, making it very valuable to use real census data, as done in our study. The incubation time, contagiousness period, and age-dependent strength of viral shedding are all important in order to predict the time profile of the viral shedding of an individual and the overall reaction time to lockdown measures.

Many ingredients are intrinsic to the population and do not change as a function of time,

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

while others do vary. An important one is the probability of being diagnosed, which was initially linked to the appearance of strong symptoms, but later extended to people with less or no symptoms due to more awareness, contact-tracing and mass testing campaigns. As a result, the probability of detection in adults and children increased, reversing the proportions of diagnosed people by age-strata. The lockdown measures were continuously adjusted by the authorities to maintain a delicate equilibrium between the contention of the virus and sustaining the economic and educational activities. Mobile phone data, clearly correlated to these measures (see **Methods**) and available separately for work activities and leisure, are used to predict the time dependence of the level of contacts. The simulation also needs to take into account the increased number of contacts taking place during the summer period in order to generate the conditions for the surge of a second wave.

In conclusion, our simulation tool leads to a deep understanding of COVID-19 dynamics in the province of Barcelona. It is able to reproduce the very different characteristics of the two waves, the underlying dynamics of the second one being much more complex. The stochastic approach allows an easy implementation of the many relevant determinants of the problem with the maximum available granularity, including time dependence. The main advantage of the tool is its capacity to study the relative impact of individual factors and to identify the most critical ones (see **Methods**). The default configuration does reproduce the relevant features of the data, giving confidence in the prediction of relative variations. More detailed calibration and validation with higher granularity data over longer period of time should bring further insight and precision. The entire machinery can be easily adapted to study the epidemiological picture of any other region for

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

which individual socio-demographic microdata are available. We can also implement the clinical history of hospitalised patients. In the constant stream of new information about the disease, further data regarding vaccination campaigns, coexistence of different variants of the virus, impact of super-spreaders, effect of local small outbreaks, etc., could be easily implemented without having to change the structure of the simulation.

Methods

The simulation described in this work reproduces the first and second wave of the COVID-19 outbreak in the province of Barcelona using both epidemiological data and population statistics, linking the "who", "where", "how" and "when" with the characteristics of the virus itself. In this section we detail the different aspects of the calculation. In particular, we discuss: the underlying compartmental model; the data used to describe the population; the description of contacts between individuals; how mobility data are processed and used to simulate the impact of lockdown measures; the implementation of other mitigation strategies (such as the use of protective masks); how the model is calibrated with the data; and finally, a description of the sensitivity of the simulation outcome to some of the most relevant parameters. The latter gives insight on the uncertainty of the prediction related to the limited knowledge of the parameters and the probabilistic nature of the problem.

SEIDR model We consider a model where the population is divided into five compartments: "susceptible", "exposed", "infected", "diagnosed" and "recovered", as shown schematically in

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Figure 6 (note that in our model we do not include traditional vital dynamics: births and deaths). When susceptible individuals enter in contact with infected persons, their state may change to exposed according to the probability

$$P = 1 - e^{-\lambda \cdot \Delta t} \tag{1}$$

where the force of infection λ is proportional to the total viral load to which the individual is exposed per unit time (day) and Δt is the time interval (1/3 day).

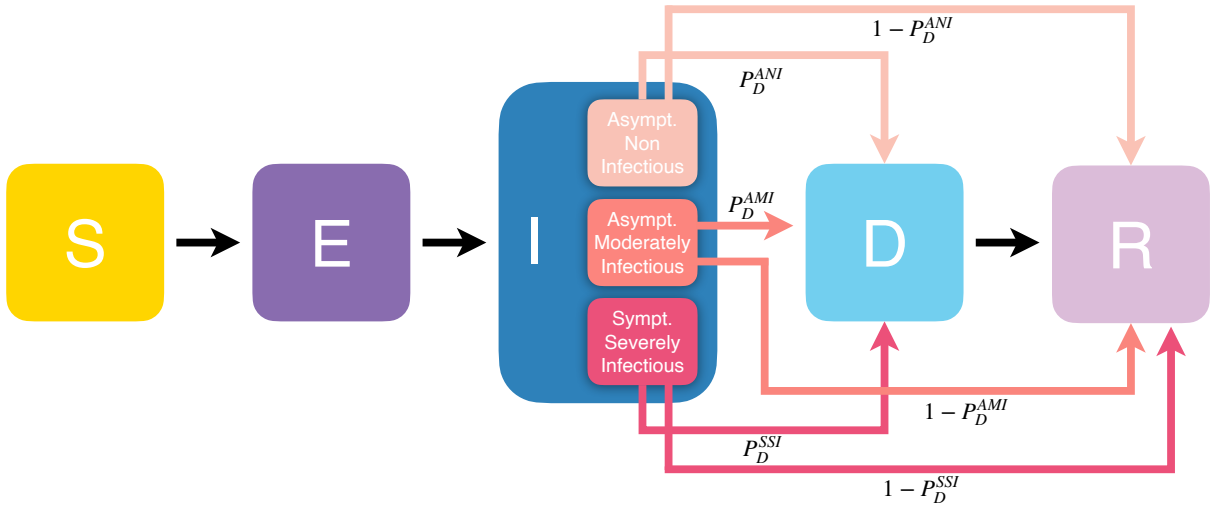


Figure 6: Transition diagram for our SEIDR model. When becoming infected, individuals are assigned a sub-compartment of disease characteristics with different symptomatology and infectiousness according to their age strata (0-14y; 15-64y; 65+y): asymptomatic non infectious (ANI), asymptomatic moderately infectious (AMI) and symptomatic strongly infectious (SSI). In turn, individuals are assigned the specific probability of being diagnosed of their sub-compartment (P_D) and conversely of transiting directly to the recovered compartment $1 - P_D$.

If exposed, individuals are assigned a given set of variables to describe the evolution of the disease in their specific case. Every exposed person will eventually transit to the infected state. Epidemiological data are available for the incubation time defined as the time between infection

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

and onset of symptoms [22]. These data are described by a Gamma distribution with average of $\mu =$ 4.58 days and standard deviation $\sigma = 3.24$ days. The most probable value for incubation time is 3 days and the maximum allocated incubation time is 14 days. Epidemiological data show that the infectiousness process (shedding of virus) starts about two days before the onset of symptoms [23]. The beginning of that process defines the transition from the susceptible to the infected state in the simulation and the beginning of the viral load shedding, whose time dependence is described by a Gamma function with average $\mu = 2.5$ days and $\sigma = 1.5$ days. The maximum emission takes place after about 1.3 days. So, people shed a significant part of their total emission of viral load before the appearance of symptoms.

Further characteristics of the disease are set in the simulation according to the results of studies of epidemiological data from the beginning of the epidemics [22, 23, 24, 25, 26, 27, 28]. Each member of the infected compartment is assigned one of three classes of symptoms and infectiousness: asymptomatic non-infectious (ANI), asymptomatic moderately infectious (AMI) and symptomatic strongly infectious (SSI) (Figure 6). The probability of belonging to such classes is strongly age-dependent: older people have a higher probability to develop symptoms and be more infectious, while children are more likely to be asymptomatic non-infectious. In addition, we assume a relation between the intensity of symptoms and infectiousness: infectious people with symptoms emit about twice as much viral load as asymptomatic infectious. In turn, each category has a specific associated probability of being diagnosed, which varied along the year. During the first wave, PCR tests were done mainly for people with strong symptoms, while during the second wave they would include close contacts of identified cases and random mass screenings. This is

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

reflected in the setting of probabilities of detection. In the first wave $P_D^{SSI} = 1$, while P_D^{ANI} and $P_D^{AMI} = 0$. In the second wave $P_D^{SSI} = 1$, while $P_D^{ANI} = 0.4$ and $P_D^{AMI} = 0.5$. As a result, the probability to be diagnosed for the age strata 65+y is 70% in the first wave and goes up to 85% in the second wave, for the age strata 15-64y the probability changes from 15% to 55%, and for the age strata 0-14y from 1% up to 30%. In the simulation, the transition takes place in eight steps between June 25 and July 7, causing some discontinuities in the daily number of newly diagnosed people as seen in Figure 2b. People diagnosed are informed of the result after a time calculated according to a Poisson distribution with an average of six days since the start of infectiousness, a minimum of three and a maximum of 14 days [27]. In the simulation diagnosed individuals are immediately put in effective quarantine and are thus no more contagious. All infected and diagnosed people will eventually move to the recovered state, at the latest 14 days after the start of infectiousness [29]. Hospitalisation, admission to intensive care and possible death have not been implemented so far. Recovered people are considered immune.

Population data To simulate the history of each of the 5.5 million citizens of the Catalan province of Barcelona, we create a directory, based on the most updated census ("Cens de població i habitatges") [15], with an entry for every individual. The original data set consists of approximately 400,000 people, each carrying an effective weight to account for the whole population. Each element is characterised by a household identifier, age, gender and further relevant characteristics (e.g., occupation, location of the workplace, commuting time - if any -, commuting means of transport - if any -, etc.). The province of Barcelona is divided into 83 areas: notably, the city of Barcelona is divided into 51 neighbourhoods, while the rest of the province is split into 32 regions

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

distributed in concentric areas around Barcelona city (see **Supplementary Material**).

Nursing facilities During the first months of the health emergency, the COVID-19 disease caused a high number of deaths among the elderly. In particular, the long-term care facilities suffered from an unfortunate notoriety due to the high mortality rate of their residents compared to people of the same age group living in family dwellings. The census we employed to reconstruct the population does not take into account people aged 65 years or over who reside in nursing facilities. In order to include this sector of the population, we used the list of the 756 official long-term care facilities established in 2019, along with details regarding the available places therein[30]. The relative age structure of the elderly population living in these facilities is almost symmetrically inverse to that of the elderly population living in family dwellings[31]: the oldest among the elderly live mainly in institutions, while the youngest live in private homes. Given the average occupancy level of nursing facilities at 86%[32] and the location of each of the facilities, we reconstruct the profile for the estimated population of 40,000 individuals hosted in nursing homes. The ages of each of the individuals are randomly assigned according to the overall age structure.

Description of activities and contacts The history of activities of every individual is simulated as a function of time. The time is organised in weeks, with five labour days and two weekend days. A day is subdivided into three 8-hour intervals. The three daily time intervals correspond to time spent: i) at home, ii) at work/school for workers and pupils (or at home for the rest), and iii) in generic *social activities*. During the weekend, people are either at home or involved in social activities. The weekly pattern of occupations, together with the granularity of 8-hour interval, is

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

303 customizable. This pattern is modified with time according to lockdown measures (e.g., closure of
304 schools or promotion of teleworking). We define five categories of contacts:

305 i Home: The home contacts include the explicit group of individuals sharing the same house-
306 hold or nursing facility. The average number of home contacts per person is 2.73.

307 ii Work and School: We assume closed groups whose sizes follow a Poisson distribution with
308 mean 10 for work, and fixed sizes for school classes varying from seven for 0-year-old to 28
309 above 12-year-old [33]. In the simulation all workers and students are assigned a company or a
310 school class, which is constant for the whole simulation period. Contacts happening in these
311 reduced groups are stable and long-lasting. In both cases, we use the census information
312 on both the proximity of the workplace and commuting time to simulate a geographical
313 distribution of companies and schools. There are 2.2M workers and 950k school pupils. On
314 average, the resulting number of contacts per person at work or at school is 13.6.

315 iii Community: Here we include social contacts (other than in the family or occupational bub-
316 bles) that we generically label as "Community" contacts. The number of assigned contacts
317 is distributed as a Gaussian with means taken from the largest and most up-to-date synthetic
318 contact matrices set [16]. The average number of contacts is age-dependent, but not the age
319 profile of these contacts. Contacts are randomly assigned within the same region and are
320 again assumed to be stable contacts but not of the "bubble" type: indeed, each element A
321 has its own fixed "Community" network $\mathscr{C}$ (the "friends circle of A"), however, an element
322 $B \in \mathscr{C}$ has a "Community" network $\mathscr{C}'$ that does not necessarily coincide with that of A. On

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

323 average, such contacts sum up to 4.8 per person.

324 iv Public transport: A fraction of people use public transportation to commute and a further
325 fraction for community activities. This may lead to additional occasional contacts beyond
326 the network of stable contacts described above. Using the most recent survey on the use of
327 means of transport in the province of Barcelona [17], we estimate that 20% of schoolchildren
328 plus 30% of workers use public transport on their way to school and workplace. We further
329 estimate that 10% of the total population uses public transport for other personal reasons.
330 We assume an effective number of about one occasional such contact per round trip and that
331 the users all carry the average viral load of the population using public transport to go to
332 work/school, or for community activities, depending on the case.

333 v Occasional contacts during community activities:

334 Finally, there may be further contacts of occasional nature, similar to the case of the public
335 transport. We assume that such contacts took place during the summer 2020 as a result of
336 the relaxation of lockdown measures coinciding with typical summer activities in crowded
337 outdoor settings (traditional festivals, cultural events, beach tourism). Some of these activi-
338 ties included also the additional presence of travellers from outside the province, modifying
339 the effective population[34]. In the simulation, the number of such contacts is set to one per
340 person with people carrying a viral load equal to the average of the full population.

341 **Mobility data** To control the spread of the disease, authorities imposed lockdown measures re-
342 stricting the mobility, starting from March 16, 2020 and lasting for the rest of the year. We use the

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

mobility data provided by the Instituto Nacional de Estadística (INE), extracted from the analysis of the position of more than 80% of the mobile phones by the three main Spanish mobile phone operators [19]. Only phones with Spanish numbers are considered in the study. The INE aggregated these data into subgroups of population that remain, enter or leave a set of "mobility areas" [20]. Data are provided for the year 2020, starting from March 16, both for weekdays and weekends. They are normalised to reference data from November 2019. We use the average data of the full Barcelona province as an indicator of work mobility and leisure mobility. The normalised intensity of mobile phones traffic is shown in Figure 7.

Lockdown measures To simulate the impact of the lockdown measures [35], we combine two methods. Measures like school closure are directly implemented by changing the configuration of the activity in the time intervals. For example, pupils stay at home instead of going to school, and do not use public transport. For the reduction of work or community contacts, we rely on the time evolution of the mobility, as justified by the correlation between mobility changes and confinement measures observed in Figure 7. There is a a substantial dip in mid-March when the first lockdown was implemented, after which the mobility slowly increases. A smaller effect is observed around mid-August (possibly related to holidays), and a more pronounced one later towards the end of October associated with a tightening of the measures. Such correlation was also observed in other regions across Europe [36]. In the simulation, we use the reduction of mobility to calculate the fraction of citizens teleworking or staying at home instead of participating to social activities.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

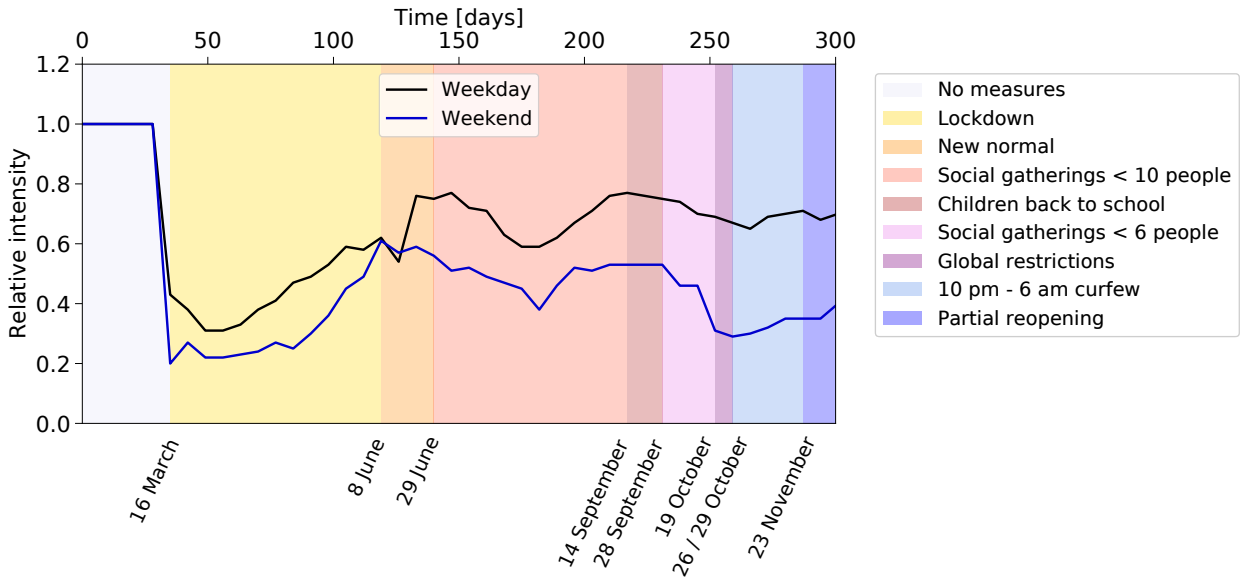


Figure 7: Relative intensity of mobile phones traffic and measures affecting work and community activities. Ratio between the intensity of mobile phones traffic during the period covered by the simulation, and for a reference week in November 2019. The pattern is influenced by the several containment measures adopted along the year. On March 16, 2020 the Spanish government decreed a state of alarm imposing of a succession of lockdown measures starting on the same day and lasting until June 8. The number of people in gatherings was limited to 10 and 6 persons on June 29 and September 26, respectively. Teleworking was encouraged, shops' and cultural events' capacity was reduced to 30% and 50%, respectively. Among other containment measures enforced on October 19, restaurants were closed, and university lectures made virtual.

Further mitigation strategies The use of protective masks was initially deterred (mostly due to the attempt of diverting the available masks to the healthcare providers), later encouraged and finally imposed. Studies show that wearing masks reduces virus transmission, although the extent of such reduction is difficult to quantify precisely [37, 38]. The effect of wearing a mask is simulated by introducing a factor reducing the viral load transmission at work and in community activities and set to roughly estimated values, 0.7 when it was encouraged and 0.35 when imposed. We assume

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

that people in nursing homes started wearing masks inside the facilities as a measure of precaution. We consider that part of the population, mostly young adults, relaxed the mask-wearing discipline towards the end of the year. In the simulation, we limit the reduction of viral load transmission from 0.35 to 0.85 for 13 to 35 years old people, from the end of October onward.

Calibration procedure The simulation counts with a total of 64 parameters to characterise the disease (26), the contacts (32) and the lockdown and self-protection measures (6) (see **Supplementary Material**). Most of them are set according to external information. Out of the 64 parameters that control the simulation, only 8 are adjusted with data as described below. First, we need to calibrate the scale factor converting the viral load calculated in arbitrary units into the force of infection leading to the correct probability of being infected. This probability governs the time needed to reach the onset of the explosive contagion regime and its growth rate. Second, for each specific value of the calibration parameter, we adjust the starting time of the lockdown measures in the simulation in order to approximately reproduce the rate of 1500 diagnosed people per day observed in the data at the peak of the first wave, as shown in Figure 3. We check that the growth rate is well reproduced: it should take about 10 days to go from 15% to 85% of the peak value. After adjusting the time, we see that the initial 50 infections happened five weeks before the start of lockdown measures, on February 10. Although the arrival time of the first infectious people in Catalonia is not precisely known, such date seems reasonable in view of the international context of the development of the disease. Third, the three categories of infectiousness need to be assigned a probability to be diagnosed for the first and the second wave. The probability of the strongly symptomatic to be diagnosed is set to one in both cases. The probability to be diagnosed for

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

asymptomatic non-infectious and infectious was initially very low and increased significantly during the second wave. These two parameters are set to zero in the first wave and have been adjusted in order to approximately reproduce the observed proportion of diagnosed people by age-categories in the second wave (Figure 3). Fourth, additional summer contacts are needed in order to increase the risk of regrowth and facilitate the appearance of a second wave. The four corresponding parameters (number of contacts, segment of population affected, and starting and ending dates) have been adjusted accordingly, in order to also respect the observed proportions of diagnosed people in the different regions (Figure 5a).

Sensitivity to Parameters

We studied the sensitivity of the simulation to the parameters and identified the most critical ones. This is relevant to understand the weakest points of the mitigation strategies. It also indicates what should be better known for an improved precision of the simulation.

First we report on the impact of parameters governing the response of individuals to the disease.

i If we reduce the overall scale of infectiousness by a factor of two, it takes twice as much time to reach the explosive contagion regime.

ii Reducing by a factor of two the probability to be diagnosed in the asymptomatic non-infectious and asymptomatic infectious categories during the second wave significantly re-

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

duces the number of people being diagnosed, and subsequently put in quarantine. As a consequence a factor of 2.5 more people become infected and the number of people diagnosed increases by 30%.

iii The time it takes for people to get diagnosed is drawn from a Poisson distribution with mean six [27]. A minimum of three days is set for the notification of the diagnosis. Removing the minimum does not have much impact, but setting it to large values does. If it is set to six days, people are put later in quarantine and more contagions can take place. The number of diagnosed people in the second wave doubles at the same time as the delay in the notification delays the peak by about a week.

iv The effectiveness of the mask is important in the surge and strength of the second wave. The simulation shows that without that protection a second wave would have quickly developed. Changing the nominal reduction of viral load transmission of 0.35 by ± 0.05 changes the number of diagnosed people by $\pm 50\%$. With much stronger reduction no second wave would develop.

Next we investigate the sensitivity to the pattern of contacts.

i The size of the contact groups at work is given by a Poisson distribution with mean 10. Changing it to 5 (20) leads to a reduction (increase) of 15% of diagnosed people. The impact is smaller in the second wave due to teleworking.

ii In addition to the 950k school pupils there are 95k university students. Since our census data

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

did not provide detailed information for this group and there were essentially no presential classes at university during 2020 after the start of the first lockdown, we considered that this group neither studied, nor worked. To analyse the impact of this assumption, we aggregated them to the pupils, assuming an average size of closed groups of contacts of 15. They resumed presential classes in mid-September. As a result, the peak of the first wave increased by 25% without readjusting the time of lockdown, while the peak of the second went down by 25%.

iii We studied the impact of making the number of community contacts dependent on the density of the population. We tested a model where the proportionality factor varies from 0.6 to 2, between Berguedà and Barcelonès, respectively, the lowest and highest density regions. The average number of community contacts increases from 4.8 to 7.6. During the first wave, the total number of diagnosed increases by 30%. The overall effect is less noticeable during the second wave, but the number of diagnoses in the Barcelonès starts dominating, unlike what is observed in Figure 5a. Such a strong density dependence is thus disfavoured.

iv The number of additional summer contacts is much more critical than the contacts in transport as they affect potentially a much more significant fraction of the population. Doubling the default level of one contact generates an earlier wave during the period where a relatively flat distribution is observed, while dividing by two reduces the level of diagnoses in early summer by a factor of six, and no second wave develops. We also tested the impact of limiting the extra contacts to the Barcelonès area rather than the whole province. However, this option does not reproduce well the sharing of diagnosed cases in the different regions

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

(Figure 5a).

v Another factor that may influence the evolution of the disease is the location of the initially infected people. In the default configuration, they are randomly distributed geographically in the province. We kept the same total number but randomly varied their location. In the first wave, this results in a variation of the onset time by ± 2.5 days. If we keep the starting date of lockdown fixed, the maximum of the first wave varies by about $\pm 30\%$. Indeed the starting time of the lockdown measures is quite critical, since one week delay typically results in a factor of three higher number of diagnosed people per day at the peak. The cases with upward fluctuations of the number of diagnosed people in the second wave are, in general, not correlated with the cases of upward fluctuations in the first wave. Only in about 1% of the cases, when the first wave developed very early, the number of infected people grew a factor of three more before the lockdown took effect (the date was kept fixed). These people eventually became immune, and later, no strong second wave was able to develop.

Finally, we studied variations in the description of mobility.

i To test the relevance of mobility and its correlation with the time dependence of the number of diagnosed people, we artificially introduced an additional four weeks of flat level of mobility in July. This has a strongly correlated effect on the time-dependent pattern of the spread, showing that mobility is a crucial ingredient for describing the evolution of the pandemic.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

ii The mobility data came in two sets with different methods and normalisations. Increasing the relative normalisation of the second set (after mid-June) by 10% almost doubles the level of infected people and shifts the peak earlier by one week.

In summary, we have explored part of the phase space of the many parameters playing a role in the expansion of the virus. The most sensitive ones include the effect of the protective masks, the probability of diagnosis, the characteristics of seasonal (summer) contacts and mobility. In some cases, more precise external information can be gathered (e.g. likelihood of diagnostic), or the full granularity of the information could be used, leading to possible improvement (e.g. mobility per region). More complex are the cases of protective masks or summer contacts. Improvements could come from a more detailed description of the parameters together with a comparison with high granularity data in terms of age and location.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Data availability

The census data used to build the complete description of the population of the Barcelona province are available from CED and the Instituto Nacional de Estadística (INE, Spanish National Institute of Statistics: 2011 Population and Housing Censuses). However, restrictions apply to the availability of these data, which were used under license for the current and previous studies, and so are not publicly available. The census data and the derived data describing the entire population of the Barcelona province are however available from the authors upon reasonable request and with permission of CED and INE. The data describing diagnosed people in 2020 in the province of Barcelona, aggregated by age or by county, are publicly available at the Agència de Qualitat i Avaluació Sanitàries de Catalunya (AQuAS) (https://aquas.gencat.cat/ca/actualitat/ultimes-dades-coronavirus/). The mobility data are publicly available from Instituto Nacional de Estadística (INE) https://www.ine.es/en/experimental/movilidad/experimental_em_en.htm.

Code availability

The code used in the simulation is available upon request.

References

1. P.G. Walker et al. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* **369, 6502**, 413–422 (2020).

2. N.G. Davies et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

demand for hospital services in the UK: a modelling study. *The Lancet* **7**, e375–e385 (2020).

3. The Computational Biology and Complex Systems group. https://biocomsc.upc.edu/en (2021).

4. Agència de Qualitat i Avaluació Sanitàries de Catalunya (AQuAS). https://aquas.gencat.cat/ca/actualitat/ultimes-dades-coronavirus/ (2021).

5. Evolution of number of cases and Rt of SARS-CoV2, Agència de Qualitat i Avaluació Sanitàries de Catalunya, (AQuAS). https://dlscitizens.blob.core.windows.net/rtreports/archived/20201231/CAT/InformeCasosRt_CAT_09.pdf (2020).

6. E. Estrada. Modeling the Present, Looking at the Future. *Physics Reports* **1869**, 1–51 (2020).

7. H. Khataee et al. Effects of social distancing on the spreading of COVID-19 inferred from mobile phone data. *Scientific Reports* **11**, 1661 (2021).

8. E. Burns et al. The natural history of symptomatic COVID-19 during the first wave in Catalonia. *Nature Communications* **12**, 777 (2021).

9. M. Castro et al., The turning point and end of an expanding epidemic cannot be precisely forecast. *PNAS* **117** (42) 26190-26196 (2020).

10. C. O. Wilke and C. T. Bergstrom, Predicting an epidemic trajectory is difficult. *PNAS* **117** (46) 28549-28551 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

514 11. A. Arenas et al. Modeling the Spatiotemporal Epidemic Spreading of COVID-19 and the Im-
515 pact of Mobility and Social Distancing Interventions. *Phys. Rev. X* **10**, 041055-1–21 (2020).

516 12. N. Ferguson et al. Strategies for containing an emerging influenza pandemic in Southeast Asia.
517 *Nature* **437**, 209–214 (2005).

518 13. M. Small and D. Cavanagh Modelling Strong Control Measures for Epidemic Propagation
519 With Networks – A COVID-19 Case Study. *IEEE Access* **8**, 109719–109731, (2020).

520 14. N.N. Chung and L.Y. Chew, Modelling Singapore COVID-19 pandemic with a SEIR multiplex
521 network model. *Sci Rep* **11**, 10122 (2021).

522 15. Instituto Nacional de Estadística (INE). Spanish National Institute of Statistics: Population
523 and Housing Censuses.
524 https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176992
525 &menu=ultiDatos&idp=1254735572981 (2011).

526 16. K. Prem et al. Projecting contact matrices in 177 geographical regions: an update and compar-
527 ison with empirical data for the COVID-19 era.
528 https://doi.org/10.1101/2020.07.22.20159772 (2020).

529 17. Enquesta de Mobilitat en Dia Feiner (EMEF) - 2019.
530 https://www.atm.cat/web/es/observatori/encuestas-de-movilidad.php (2019).

531 18. T. Moreno et al. Tracing surface and airborne SARS-CoV-2 RNA inside public buses and
532 subway trains. *Environment International* **147**, 106326 (2021).

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

19. Instituto Nacional de Estadística (INE). Studies on mobility based on mobile phone https://www.ine.es/en/experimental/movilidad/experimental_em_en.htm (2020).

20. Instituto Nacional de Estadística (INE). Functional Urban Areas: Barcelona https://ine.es/index.htm (2021).

21. J.M. Arauzo-Carod et al. Do local characteristics act in a similar way for the first two waves of COVID-19? Analysis at intraurban level in Barcelona. *J Public Health (Oxford)* **43,3**, 455-461 (2021).

22. Q. Bi et al. Epidemiology and Transmission of COVID-19 in 391 Cases and 1286 of Their Close Contacts in Shenzhen, China: A Retrospective Cohort Study. *The Lancet. Infectious Diseases* **20,8**, 911–919 (2020).

23. X. He et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* **26**, 672–675 (2020).

24. L. Di Domenico et al. Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. *BMC Med* **18**, 240 (2020).

25. A. Aleta et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Social Behaviour* **4**, 964-971 (2020).

26. R. Verity et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet. Infectious Diseases* **20,6**, 669-677 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

27. Análisis de los casos de COVID-19 notificados a la RENAVE hasta el 10 de mayo en España. Informe COVID-19 nº 33. 29 de mayo de 2020. https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx (2020).

28. A. Soriano-Arandes et al. Household Severe Acute Respiratory Syndrome Coronavirus 2 Transmission and Children: A Network Prospective Study. *Clinical Infectious Diseases* **73,6**, e1261–e1269 (2021).

29. T. Tolossa et al. Time to recovery from COVID-19 and its predictors among patients admitted to treatment center of Wollega University Referral Hospital (WURH), Western Ethiopia: Survival analysis of retrospective cohort study. *PLOS ONE* **16 (6)**, e0252389 (2021).

30. Centro de Ciencias Humanas y Sociales, CSIC. Envejecimiento en Red, datos de abril de 2019. http://envejecimiento.csic.es/documentos/documentos/enred-estadisticasresidencias2019.pdf (2019).

31. Instituto Nacional de Estadística (INE). Una estimación de la población que vive en residencias de mayores. http://envejecimientoenred.es/una-estimacion-de-la-poblacion-que-vive-en-residencias-de-mayores/ (2020).

32. Diputació de Barcelona. Informació Estadística Local. https://www.diba.cat/hg2/presentacioprov.asp?prid=954 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

571 33. Departament d'Ensenyament - Generalitat de Catalunya. Ràtios d'alumnes per estudi i unitat
572 o grup.
573 https://educacio.gencat.cat/ca/departament/estadistiques/indicadors/sistema-
574 educatiu/escolaritzacio/ratios/ (2020).

575 34. Departament d'Estadística i Difusió de Dades - Ajuntament de Barcelona. Demanda hotelera
576 a Barcelona. Serie històrica 2005-2021.
577 https://ajuntament.barcelona.cat/estadistica/catala/Estadistiques_per_temes/
578 Turisme_i_promocio_economica/Turisme/Oferta_demanda_hotelera/evo/th07.htm (2021).

579 35. Diari Oficial de la Generalitat de Catalunya.
580 https://dogc.gencat.cat/ca/inici/ (2020).

581 36. S. Santamaria et al. Measuring the impact of COVID-19 confinement measures on human mo-
582 bility using mobile positioning data. A European regional analysis. *Safety Science* **132**, 104925
583 (2020).

584 37. Y. Cheng et al. Face masks effectively limit the probability of SARS-CoV-2 transmission.
585 *Science* **372, 6549**, 1439-1443 (2021).

586 38. Y. Wang et al. How effective is a mask in preventing COVID-19 infection? *Medical devices &*
587 *sensors* **4**, e10163 (2021).

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Acknowledgments

The authors affiliated to CED, IFAE and i2CAT acknowledge the support of the CERCA institution, Centres de Recerca de Catalunya.

MB, XJ, LLM, MMar, PM, JP, IR acknowledge support from the grant 2020PANDE0180 of the programme PANDEMIES 2020, "Replegar-se per créixer: l'impacte de les pandèmies en un món sense fronteres visibles" of the Agència de Gestió d'Ajuts Universitaris i de Recerca of the Generalitat de Catalunya. LG thanks the funding from the European Union's Horizon 2020 research and innovation programme under grant agreement ID 758145. AL acknowledges the support from the Talent Research Program (Universitat Autònoma de Barcelona). PM has received funding from the Spanish Ministry of Science and Innovation (PID2020-112965GB-I00/AEI/ 10.13039/501100011033), from the Agency for Management of University and Research Grants of the Government of Catalonia (project SGR 1069), and also received support from Ajuntament de Barcelona. MMan acknowledges support from Marie Sklodowska-Curie grant agreement ID 6655919. VV has been partially supported by the H2020 programme and by the Secretary of Universities and Research of the Government of Catalonia through a Marie Skłodowska-Curie COFUND fellowship – Beatriu de Pinós programme ID 801370.

Author Contributions

MB led the conception of the project. MB, LLM, MMan, PM, VV contributed to the modelling design. MB, LG, XJ, AL, LLM, IR took care of data preparation. MB, LG, LLM, MMan, PM,

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

AP, IR developed the code. MB, LLM, PM, VV wrote the manuscript. All authors contributed to the discussion and interpretation of the results, revised critically the draft and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Supplementary material.

M. Bosman[1*], A. Esteve[2,3], L. Gabbanelli[1], X. Jordan[4], A. López-Gay[2,5], M. Manera[1,6], M. Martínez[1,7], P. Masjuan[1,8], Ll.M. Mir[1], J. Paradells[4,9], A. Pignatelli[1], I. Riu[1], V. Vitagliano[1,10,11]

[1] *Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Barcelona, Spain*

[2] *Centre d'Estudis Demogràfics (CED), Barcelona, Spain*

[3] *Serra Húnter Fellow, Departament de Ciències Polítiques i Socials, Universitat Pompeu Fabra, Barcelona, Spain*

[4] *i2CAT Foundation, Edifici Nexus (Campus Nord UPC), Barcelona, Spain*

[5] *Departament de Geografia, Universitat Autònoma de Barcelona, Bellaterra, Spain*

[6] *Serra Húnter Fellow, Departament de Física, Universitat Autònoma de Barcelona, Bellaterra, Spain*

[7] *Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain*

[8] *Departament de Física, Universitat Autònoma de Barcelona, Bellaterra, Spain*

[9] *Departament d'Enginyeria Telemàtica, Universitat Politècnica de Catalunya, Barcelona, Spain*

[10] *Department of Mathematics and Physics, University of Hull, Kingston upon Hull, HU6 7RX, UK*

[11] *DIME, University of Genova, Via all'Opera Pia 15, 16145 Genova, Italy*

**Corresponding author: bosman@ifae.es*

Parameters of the model

The model to simulate the spread of Covid-19 counts with a total of 64 parameters to characterize the disease (26), the contacts (32) and the lockdown and self-protection measures (6). Most of them

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

are set according to external information. Out of the 64 parameters that control the simulation, only 8 are adjusted. Tables 1 to 7 describe the parameters, their settings and related references.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Table 1: Parameters of the model (part I).

Category	Number	Parameters	Source	Adjusted
		Disease		
Incubation time (Gamma distribution)	2	μ, σ = 4.58, 3.24 days	Medical data [1, 2, 3]	No
Infectiousness time (Gamma distribution)	2	μ, σ = 2. 5, 1.7 days	Medical data [1, 2, 3]	No
Onset of infectiousness in relation to onset of symptoms	1	-2 days	Medical data [2]	No
Relative viral strength of asymptomatic, moderate and severely infectiousness	3	0, 1, 2	Medical data [3]	No
Viral strength to probability of infection conversion	1	4	Adjusted	Shape & size of 1st wave
Fraction in categories of viral strength (age dependent)	2×3	See table 3	Medical data [3]	No
Probability to be diagnosed in first (second) wave	3×2	P_D^{ANI} = 0 (0.4) P_D^{AMI} = 0 (0.5) P_D^{SSI} = 1 (1)	Medical data [4]	Yes for P_D^{ANI} & P_D^{AMI} during 2nd wave
Diagnosis time (Poisson distribution)	3	μ, minimum, maximum = 6, 3, 14 days	Medical data [4]	No
Recovery time	1	14 days	Medical data [5]	No
Initially exposed	1	0.001%	Estimated	No
		Contact's model		
Workplace size (Poisson distribution)	1	$\mu = 10$	Estimated	No
Classroom size (fixed, age dependent)	6	See table 4	Departament d'Educació, Gen-Cat [6]	No
Social contacts (Poisson distribution, age dependent)	16	See table 5	Synthetic contact matrices [7]	No
Spurious summer contacts	4	See table 6	Ajuntament de Barcelona [8]	Shape of 2nd wave

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Table 2: Parameters of the model (part II).

Category	Number	Parameters	Source	Adjusted
Use of public transportation				
Fraction of workers (Work)	1	30%	Autoritat del Transport Metropolità (ATM) [9]	No
Fraction of pupils (School)	1	20%	ATM [9]	No
Fraction of population (social activities)	1	10%	ATM [9]	No
Average number of contacts per round trip (work/school) (fixed number)	1	1.0	Estimated	No
Average number of contacts per round trip (social activities) (fixed number)	1	1.1	Estimated	No
Confinement				
Date of 1st day of simulation	1	34 days before March 16, 2020 starting date of first confinement	Diari Official de la Generalitat de Catalunya [10]	Shape of 1st wave
Effect of mask	4	See table 7	Medical data [11, 12]	No
Inter-calibration of mobile data	1	1	Estudios de movilidad a partir de la telefonía móvil [13]	No

Table 3: Fraction of asymptomatic, moderately and severely infectious people depending on the age.

Age (years)	Asymptomatic	Moderately infectious	Severely infectious
< 15	0.69	0.30	0.01
$15 - 64$	0.40	0.45	0.15
> 64	0.05	0.25	0.70

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Table 4: Number of children in each age group.

Age	Classrooms' size
0	7
1	12
2	16
3 – 5	22
6 – 11	25
12 - 18	28

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Table 5: Average number of stable social contacts depending on the age.

Age	Mean
< 5	3.21
5 – 10	3.88
10 – 15	5.27
15 – 20	5.97
20 – 25	5.26
25 – 30	4.33
30 – 35	4.35
35 – 40	5.51
40 – 45	6.48
45 – 50	4.94
50 – 55	5.26
55 – 60	5.26
60 – 65	5.06
65 – 70	4.05
70 – 75	4.89
> 75	2.95

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Table 6: Parameters used to model additional summer contacts.

Parameters	Value
Average number of additional contacts	1.0
Starting date	June 8, 2020
Ending date	September 28, 2020
Fraction of population involved	100%

Table 7: The effect of wearing a mask is simulated by introducing a factor that reduces the viral load transmission at work or school and in community activities. As of October 26, 2020, a relaxation of the discipline in the use of the masks by young adults is assumed.

Use of mask	Viral load reduction
Encouraged	0.70
Imposed	0.35
Relaxed use (from 13 to 35 year-old)	0.85
Starting date of relaxation	October 26, 2020

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Statistics of the sample

Tables 8, 9 and 10 present the statistics of different subsets of the population.

Table 8: Number of people in each age category.

Age (years)	Persons
< 15	861,266
15 – 64	3,704,595
> 64	961,372
$\rightarrow$ in nursing homes	39,373
Total	5,527,233

Table 9: People working or going to school.

Occupation	Persons
Workers	2,247,353
Schoolchildren (0 to 18 year-old)	953,019
Other	2,326,861
Total	5,527,233

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Table 10: Workplaces model assumed using the information available in the census file.

Work location	Home region	Time to work	Assigned work region	Workers
At home	Any	–	Home region	197,827
Own	BCN	$<$ 20’	Home neighbourhood	190,316
municipality		$>$ 20’	Another BCN neigh.	254,219
	Counties	Any	Home region	448,229
Another	BCN	Any	Counties	128,165
municipality	Counties	$<$ 45’	Home region	659,138
		$>$ 45’	Outside home region	175,293
Several	BCN	Any	Counties	42,032
municipalities	Counties	Any	Home region	144,856
Abroad	Any	Any	Outside home region	7,278

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Description of the province of Barcelona

In the simulation, the province of Barcelona is divided into 83 regions: notably, the city of Barcelona is divided into 51 neighbourhoods, while the rest of the province is split into 32 regions. Table 11 shows the number of inhabitants in the 32 regions. These regions are distributed within three concentric areas with the city of Barcelona at the center, as shown in Figure 1: the Barcelonès – the county including the cities of Barcelona, Badalona, L'Hospitalet de Llobregat, Sant Adrià de Besòs and Santa Coloma de Gramanet, for a total of 2.3 million inhabitants –, a first "ring", which includes Baix Llobregat, Vallès Occidental, Vallès Oriental, and Maresme – 2.5 million inhabitants –, and a second "ring", including Garraf, Alt Penedès, Anoia, Bages, Berguedà, and Osona – 0.8 million inhabitants –. Table 12 shows the number of inhabitants in the 51 neighbourhoods of Barcelona, for a total of 1.6 million people.

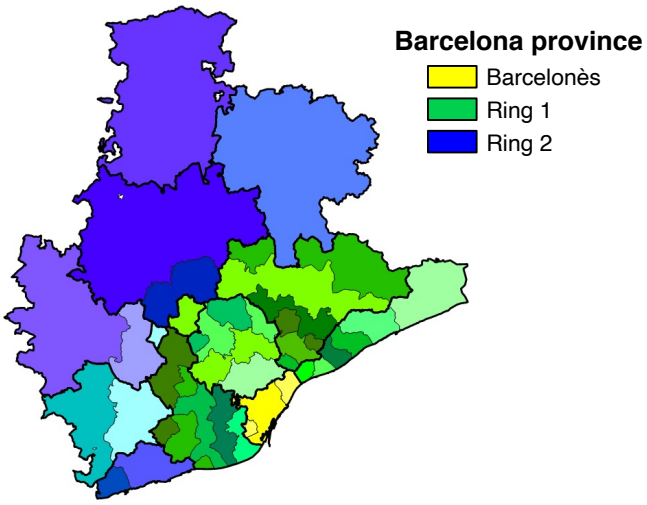


Figure 1: Map of the province of Barcelona. The regions used in the simulation regrouped in three concentric areas.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Table 11: Population in the regions defined in the province of Barcelona census file according to the county they belong to and their distance to the city of Barcelona.

Comarca	Distance (km)	Population
Alt Penedès	30 – 40	72,309
	$>$ 40	25,692
Anoia	30 – 40	29,429
	$>$ 40	88,592
Bages	30 – 40	22,174
	$>$ 40	161,985
Baix Llobregat	$<$ 10	242,777
	10 – 15	253,968
	15 – 20	118,991
	20 – 25	111,045
Baix Llobregat / Alt Penedès	25 – 30	74,370
Barcelonès (excluding the city of Barcelona)	$<$ 10	627,672
Berguedà	$>$ 65	40,292
Garraf	30 – 40	64,702
	$>$ 40	79,947
Maresme	10 – 15	28,159
	15 – 20	66,158
	20 – 25	40,793
	25 – 30	134,370
	30 – 40	46,859
	$>$ 40	115,155
Osona	$>$ 40	151,652
Vallès Occidental	10 – 15	258,591
	15 – 20	344,845
	20 – 25	238,001
	25 – 30	39,125
Vallès Occidental / Oriental	30 – 40	106,489
Vallès Oriental	10 – 15	21,748
	15 – 20	103,441
	20 – 25	25,945
	25 – 30	109,488
	$>$ 40	41,054

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Table 12: Population of the neighbourhoods of the city of Barcelona.

Neighbourhoods	Population	Neighbourhoods	Population
Canyelles, les Roquetes	20,748	el Guinardó	35,853
Diagonal i el Front Marítim del Poblenou, Provençals de Poblenou	33,018	el Parc i la Llacuna del Poblenou, la Vila Olímpica del Poblenou	23,576
Hostafrancs, la Bordeta	34,130	el Poble Sec	37,541
Les Corts	46,844	el Poblenou	33,263
Porta	25,265	el Putxet i el Farró	30,337
Sant Andreu (Est)	24,169	el Raval (Nord)	21,234
Sant Andreu (Oest)	34,272	el Raval (Sud)	22,079
Sant Antoni	37,735	el Turó de la Peira, Can Peguera, la Guineueta	34,687
Sant Gervasi-Galvany	43,750	l'Antiga Esquerra de l'Eixample	39,422
Sant Martí de Provençals	25,758	la Barceloneta	17,928
Sant Pere, Santa Caterina i la Ribera	22,348	la Dreta de l'Eixample	41,540
Sants	44,691	la Font d'en Fargues, Horta	36,028
Sants - Badal	24,277	la Marina del Prat Vermell, la Marina del Port, la Font de la Guatlla	41,664
Vallcarca i els Penitents, el Coll, la Salut	36,176	la Maternitat i Sant Ramon, Pedralbes	33,837
Vallvidrera, el Tibidabo i les Planes, Sarrià	27,762	la Nova Esquerra de l'Eixample (Nord)	29,198
Verdun, la Prosperitat	40,707	la Nova Esquerra de l'Eixample (Sud)	32,287
Vilapicina i la Torre Llobeta	26,544	la Sagrada Familia (Nord)	22,534
el Baix Guinardó, Can Baró	36,252	la Sagrada Familia (Sud)	29,771
el Barri Gòtic	17,566	la Sagrera	12,104
el Besòs i el Maresme	19,988	la Teixonera, Sant Genís dels Agudells, Montbau, la Vall d'Hebron, la Clota	29,230
el Camp d'en Grassot i Gràcia Nova	34,582	la Trinitat Nova, Torre Baró, Ciutat Meridiana, Vallbona	18,454
el Camp de l'Arpa del Clot	38,602	la Trinitat Vella, Baró de Viver, el Bon Pastor	21,213
el Carmel	30,183	la Verneda i la Pau	28,426
el Clot	28,550	la vila de Gràcia	49,576
el Congrés i els Indians, Navas	37,642	les Tres Torres, Sant Gervasi - la Bonanova	41,116
el Fort Pienc	30,706		

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

34 ## Age distribution of various sectors of the population

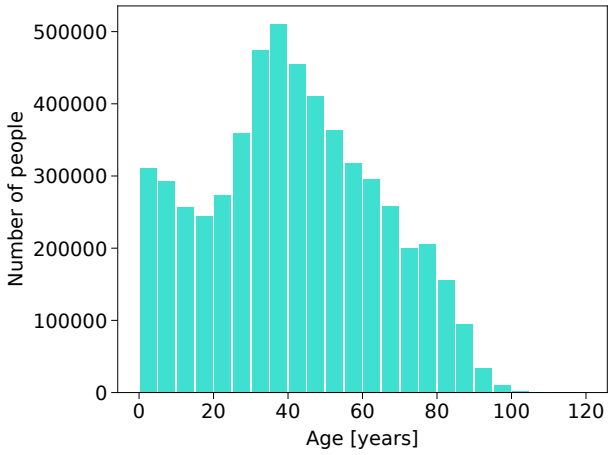


Figure 2: Population age. The age of the population of the province of Barcelona (5.5 million inhabitants) is shown.

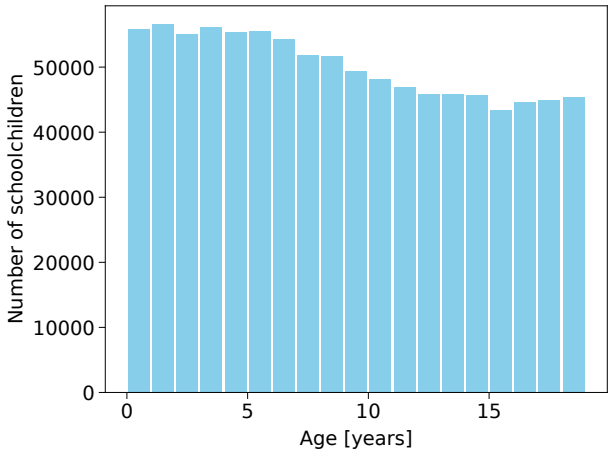
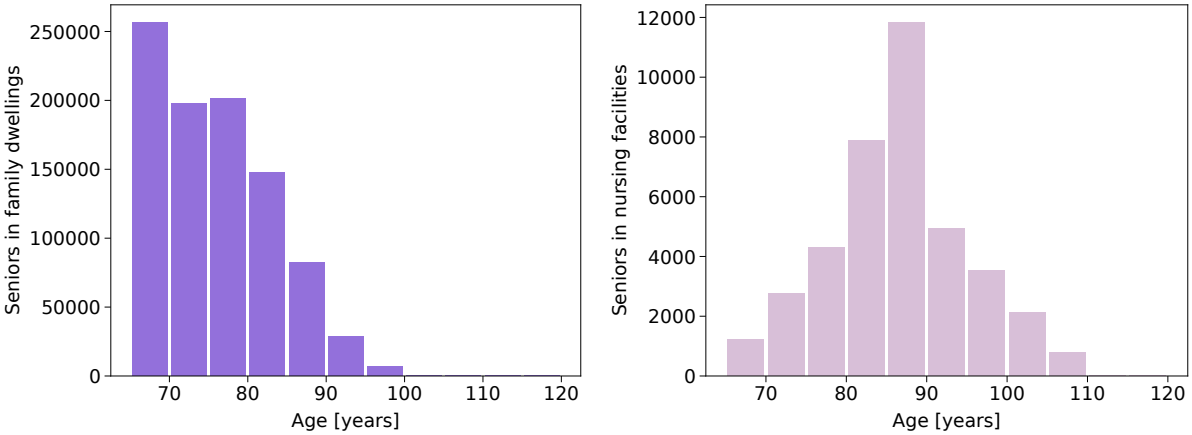


Figure 3: Schoolchildren age. The age of the 0 to 18-year-old children who attend school is shown.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .



(a) (b)

Figure 4: Seniors age. The age of people over 64 years of age is shown separately for **a** those who live in family dwellings, and **b** those who live in nursing facilities.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

35 # Time profile of incubation time and viral shedding

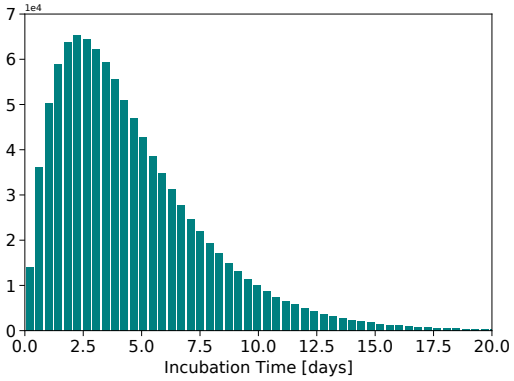


Figure 5: Incubation time. The time of incubation of the virus (in days) follows a Gamma distribution of mean 4.58 days and standard deviation 3.24 days.

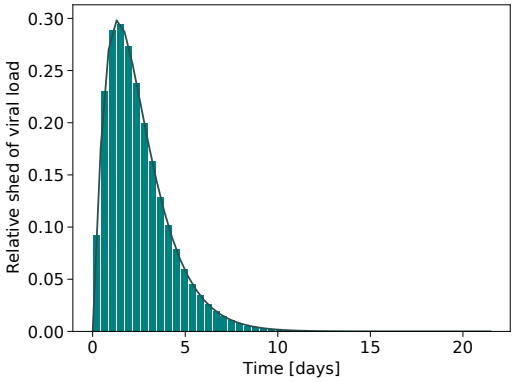


Figure 6: Relative profile of viral shedding. The distribution of the relative profile of the viral shedding follows a Gamma distribution of mean 2.5 days and standard deviation 1.7 days. In the simulation the distribution is truncated at 14 days.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

1. Q. Bi et al. Epidemiology and Transmission of COVID-19 in 391 Cases and 1286 of Their Close Contacts in Shenzhen, China: A Retrospective Cohort Study. *The Lancet. Infectious Diseases* **20.8**, 911–919 (2020).

2. X. He et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* **26**, 672–675 (2020).

3. L. Di Domenico et al. Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies *BMC Med* **18**, 240 (2020).

4. Análisis de los casos de COVID-19 notificados a la RENAVE hasta el 10 de mayo en España. Informe COVID-19 nº 33. 29 de mayo de 2020. https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx (2020).

5. T. Tolossa et al. Time to recovery from COVID-19 and its predictors among patients admitted to treatment center of Wollega University Referral Hospital (WURH), Western Ethiopia: Survival analysis of retrospective cohort study. *PLOS ONE* **16 (6)**, e0252389 (2021).

6. Departament d'Esenyament - Generalitat de Catalunya. Ràtios d'alumnes per estudi i unitat o grup. https://educacio.gencat.cat/ca/departament/estadistiques/indicadors/sistema-educatiu/escolaritzacio/ratios/ (2020).

medRxiv preprint doi: https://doi.org/10.1101/2021.12.02.21266468; this version posted December 8, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

7. K. Prem et al. Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the COVID-19 era. https://doi.org/10.1101/2020.07.22.20159772 (2020).

8. Departament d'Estadística i Difusió de Dades - Ajuntament de Barcelona. Demanda hotelera a Barcelona. Serie històrica 2005-2021. https://ajuntament.barcelona.cat/estadistica/catala/Estadistiques_per_temes/Turisme_i_promocio_economica/Turisme/Oferta_demanda_hotelera/evo/th07.htm (2021).

9. Enquesta de Mobilitat en Dia Feiner (EMEF) - 2019. https://www.atm.cat/web/es/observatori/encuestas-de-movilidad.php (2019).

10. Diari Oficial de la Generalitat de Catalunya . https://dogc.gencat.cat/ca/inici/ (2020).

11. Y. Cheng et al. Face masks effectively limit the probability of SARS-CoV-2 transmission *Science* **372, 6549** 1439-1443 (2021).

12. Y. Wang et al. How effective is a mask in preventing COVID-19 infection? *Medical devices & sensors* **4** e10163 (2021).

13. Estudios de movilidad a partir de la telefonía móvil. https://www.ine.es/experimental/movilidad/experimental_em.htm (2020).