

1 Stochastic simulation of successive waves of COVID-19 in 2 the province of Barcelona.

3 M. Bosman^{1*}, A. Esteve^{2,3}, L. Gabbanelli¹, X. Jordan⁴, A. López-Gay^{2,5}, M. Manera^{1,6}, M.
4 Martínez^{1,7}, P. Masjuan^{1,8}, Ll.M. Mir¹, J. Paradells^{4,9}, A. Pignatelli¹, I. Riu¹, V. Vitagliano^{1,10,11}

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

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

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

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

9 ⁵*Departament de Geografia, Universitat Autònoma de Barcelona, Bellaterra, Spain*

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

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

12 ⁸*Departament de Física, Universitat Autònoma de Barcelona, Bellaterra, Spain*

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

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

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

16 **Corresponding author: bosman@ifae.es*

17 **Abstract**

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

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

29 **Introduction**

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

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

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

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

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

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

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

- 86 i Socio-demographic individual microdata from the latest census data provided by the Spanish
87 National Statistics Institute ¹⁵.
- 88 ii Contacts matrices for densely populated (and country-dependent) environments ^{16, 17, 18}.

89 iii Mobility data, reconstructed from information supplied by mobile network operators ^{19, 20}.

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

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

102 **Results**

103 Our goal is to simulate the COVID-19 outbreak within the population of the province of Barcelona
104 (5.5 million inhabitants) and reproduce the first two epidemic waves experienced in the region
105 during February-June and July-December 2020 for a total of 300 days. We assume the existence of
106 an initial set of about 50 exposed people (0.001% of the population), randomly distributed across
107 the region. The actual location of people changes throughout the day and throughout the week: the

108 weekdays of workers and students, for example, are divided into three 8-hour fractions according
109 to a given pattern (home \Rightarrow job (or school) \Rightarrow other activities). The pattern of allowed locations
110 and activities evolves according to lockdown and mobility restrictions promulgated during such
111 300 days.

112 The number of individuals in each epidemiological compartment predicted by the simulation
113 as a function of time is shown in Figure 1. Without lockdown measures (Figure 1a), more than
114 90% of the total population of Barcelona and its province get affected by the disease in about 100
115 days. Taking into account lockdown measures (Figure 1b), the amount of infected people reduces
116 to 300k in about the same period. In the second wave, continuing until the end of the year, the total
117 number of infected people increases to 750k. Figure 2 presents the numbers of currently infected
118 and newly diagnosed people predicted by the simulation as a function of time. The figure presents
119 cumulative and age-dependent evolution curves. The subset of hosts of collective accommodation
120 institutions (nursing homes) is shown separately. The cumulative curve shows a first wave char-
121 acterised by a steep onset, interrupted by the strict lockdown measures established on March 16,
122 2020. The mobility restrictions induce the turn-over of the curve and its subsequent decay over a
123 time scale related to the duration of the incubation and infectious phases of the disease. Contacts
124 start increasing again during summer, leading to a second wave in autumn. Adults constitute the
125 most relevant fraction of the infected compartment during both waves. The category of diagnosed
126 is instead dominated by seniors in the first wave and adults in the second.

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

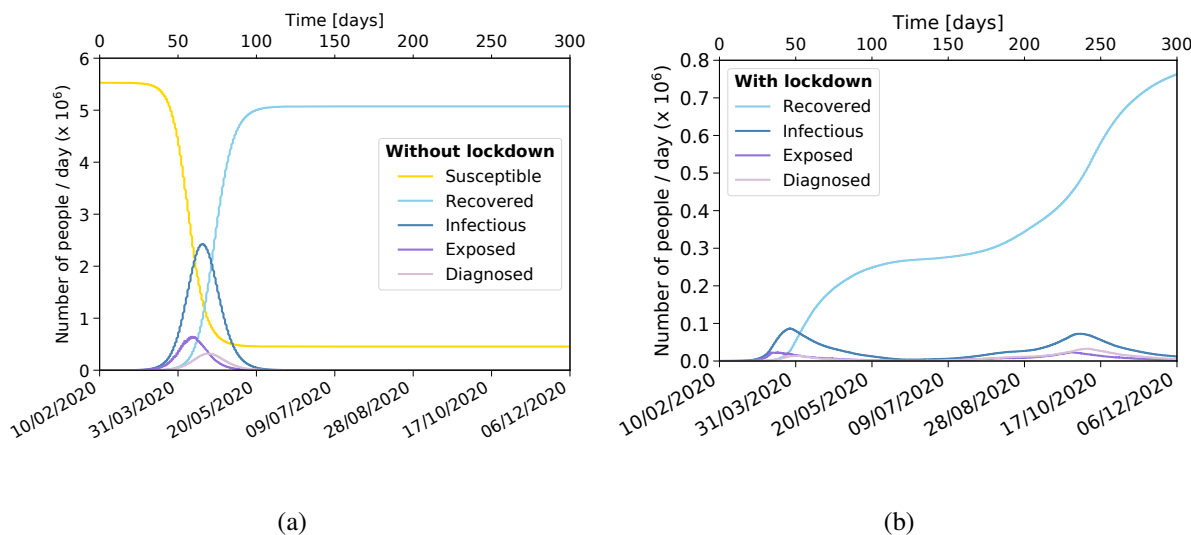


Figure 1: Evolution of the COVID-19 pandemic in the province of Barcelona. Number of susceptible, exposed, infected, diagnosed and recovered people as a function of time over a period of 300 days from February 10 to December 6, 2020. It has been assumed that 50 people, chosen at random, are exposed at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

128 region-dependence:

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

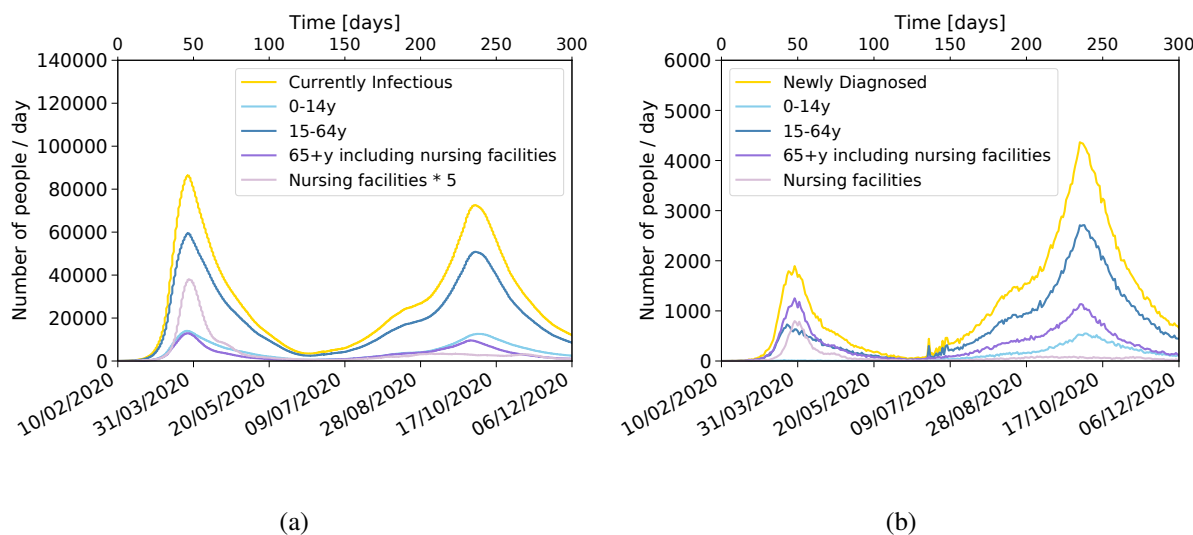


Figure 2: Time evolution of infected people. The number of infected people over a period of 300 days is shown split by age categories: less than 15 years old, from 15 to 64 years old and 65 years old and above (hosts in nursing facilities are shown separately). It has been assumed that 50 people, chosen 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.

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

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

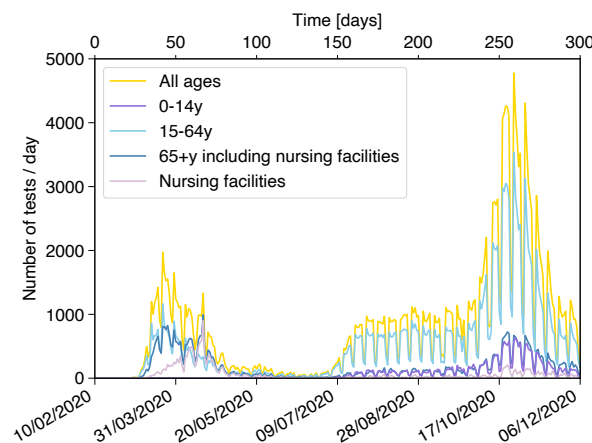


Figure 3: Number of positive PCR tests in the province of Barcelona. The number of positive PCR tests in the province of Barcelona for the period between February 10 and December 6, 2020 is shown separately for the different age segments of the population, including its sum.

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

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

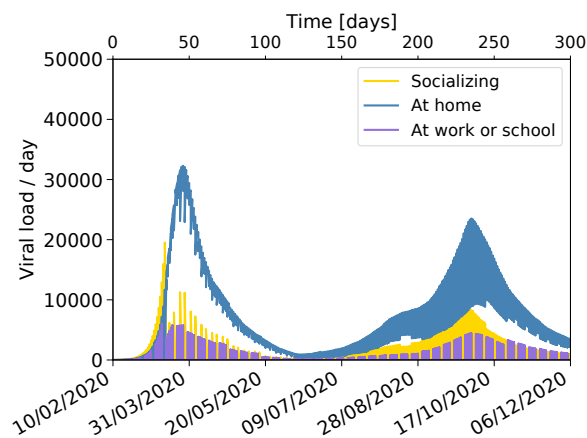


Figure 4: Total viral load evolution. The evolution of the main components of the viral load to which the population is exposed is shown separately for time spent at home, at work or school, or in social/community interaction. It has been assumed that 50 people, chosen at random, are exposed on February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

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

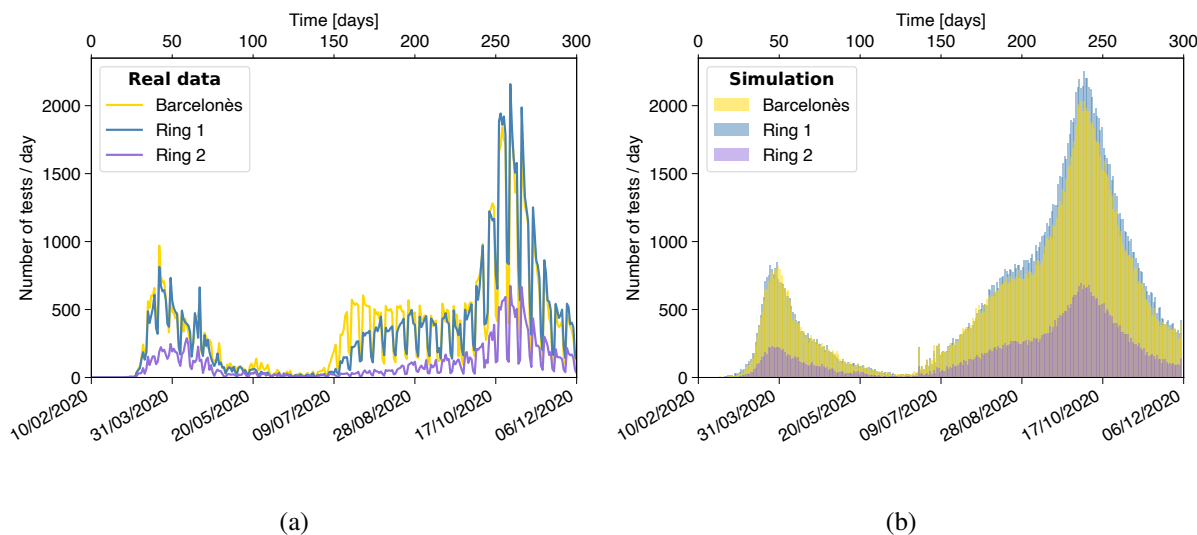


Figure 5: Number of positive PCR tests in the province of Barcelona. The results are shown separately for the Barcelonès county, and the counties surrounding it for the first and the second ring. **a** Real data. **b** Simulation.

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

178 Discussion

179 The first and second waves are of quite different nature²¹. The first one is the result of the initially
180 free propagation of the virus, later controlled by the application of strict lockdown measures. Di-
181 agnostics were strongly related to the appearance of symptoms. During the second wave, mobility
182 restrictions were alternatively relaxed and tightened depending on the epidemiological situation.
183 The use of masks proved to be essential, as well as the increase of the efficiency of diagnosis of
184 asymptomatic individuals and the corresponding quarantine measures.

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

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

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

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

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

224 **Methods**

225 The simulation described in this work reproduces the first and second wave of the COVID-19
226 outbreak in the province of Barcelona using both epidemiological data and population statistics,
227 linking the “who”, “where”, “how” and “when” with the characteristics of the virus itself. In this
228 section we detail the different aspects of the calculation. In particular, we discuss: the underlying
229 compartmental model; the data used to describe the population; the description of contacts be-
230 tween individuals; how mobility data are processed and used to simulate the impact of lockdown
231 measures; the implementation of other mitigation strategies (such as the use of protective masks);
232 how the model is calibrated with the data; and finally, a description of the sensitivity of the simu-
233 lation outcome to some of the most relevant parameters. The latter gives insight on the uncertainty
234 of the prediction related to the limited knowledge of the parameters and the probabilistic nature of
235 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

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

$$P = 1 - e^{-\lambda \cdot \Delta t} \quad (1)$$

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

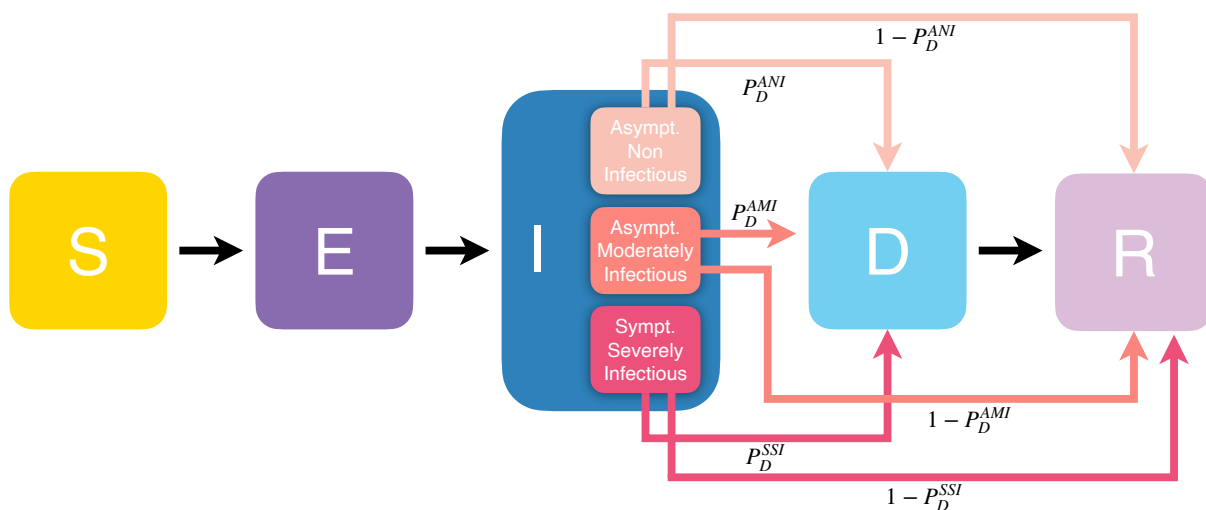


Figure 6: Transition diagram for our SEIDR model. When becoming infected, individuals are assigned a sub-compartment of disease characteristics with different symptomatology and infectiousness according to their age strata (0-14y; 15-64y; 65+y): asymptomatic non infectious (ANI), asymptomatic moderately 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 $1 - P_D$.

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

241 and onset of symptoms²². These data are described by a Gamma distribution with average of $\mu =$
242 4.58 days and standard deviation $\sigma = 3.24$ days. The most probable value for incubation time is 3
243 days and the maximum allocated incubation time is 14 days. Epidemiological data show that the
244 infectiousness process (shedding of virus) starts about two days before the onset of symptoms²³.
245 The beginning of that process defines the transition from the susceptible to the infected state in the
246 simulation and the beginning of the viral load shedding, whose time dependence is described by
247 a Gamma function with average $\mu = 2.5$ days and $\sigma = 1.5$ days. The maximum emission takes
248 place after about 1.3 days. So, people shed a significant part of their total emission of viral load
249 before the appearance of symptoms.

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

262 reflected in the setting of probabilities of detection. In the first wave $P_D^{SSI} = 1$, while P_D^{ANI} and P_D^{AMI}
263 = 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
264 to be diagnosed for the age strata 65+y is 70% in the first wave and goes up to 85% in the second
265 wave, for the age strata 15-64y the probability changes from 15% to 55%, and for the age strata
266 0-14y from 1% up to 30%. In the simulation, the transition takes place in eight steps between
267 June 25 and July 7, causing some discontinuities in the daily number of newly diagnosed people
268 as seen in Figure 2b. People diagnosed are informed of the result after a time calculated according
269 to a Poisson distribution with an average of six days since the start of infectiousness, a minimum
270 of three and a maximum of 14 days²⁷. In the simulation diagnosed individuals are immediately
271 put in effective quarantine and are thus no more contagious. All infected and diagnosed people
272 will eventually move to the recovered state, at the latest 14 days after the start of infectiousness²⁹.
273 Hospitalisation, admission to intensive care and possible death have not been implemented so far.
274 Recovered people are considered immune.

275 **Population data** To simulate the history of each of the 5.5 million citizens of the Catalan province
276 of Barcelona, we create a directory, based on the most updated census (“Cens de població i habi-
277 tatges”)¹⁵, with an entry for every individual. The original data set consists of approximately
278 400,000 people, each carrying an effective weight to account for the whole population. Each el-
279 ement is characterised by a household identifier, age, gender and further relevant characteristics
280 (e.g., occupation, location of the workplace, commuting time - if any -, commuting means of
281 transport - if any -, etc.). The province of Barcelona is divided into 83 areas: notably, the city of
282 Barcelona is divided into 51 neighbourhoods, while the rest of the province is split into 32 regions

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

284 **Nursing facilities** During the first months of the health emergency, the COVID-19 disease caused
285 a high number of deaths among the elderly. In particular, the long-term care facilities suffered from
286 an unfortunate notoriety due to the high mortality rate of their residents compared to people of the
287 same age group living in family dwellings. The census we employed to reconstruct the population
288 does not take into account people aged 65 years or over who reside in nursing facilities. In order
289 to include this sector of the population, we used the list of the 756 official long-term care facilities
290 established in 2019, along with details regarding the available places therein³⁰. The relative age
291 structure of the elderly population living in these facilities is almost symmetrically inverse to that
292 of the elderly population living in family dwellings³¹: the oldest among the elderly live mainly
293 in institutions, while the youngest live in private homes. Given the average occupancy level of
294 nursing facilities at 86%³² and the location of each of the facilities, we reconstruct the profile for
295 the estimated population of 40,000 individuals hosted in nursing homes. The ages of each of the
296 individuals are randomly assigned according to the overall age structure.

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

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-old³³. 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 set¹⁶. 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 \mathcal{C} (the “friends circle of A ”), however, an element
322 $B \in \mathcal{C}$ has a “Community” network \mathcal{C}' that does not necessarily coincide with that of A . On

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 Barcelona¹⁷, 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 population³⁴. 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

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

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

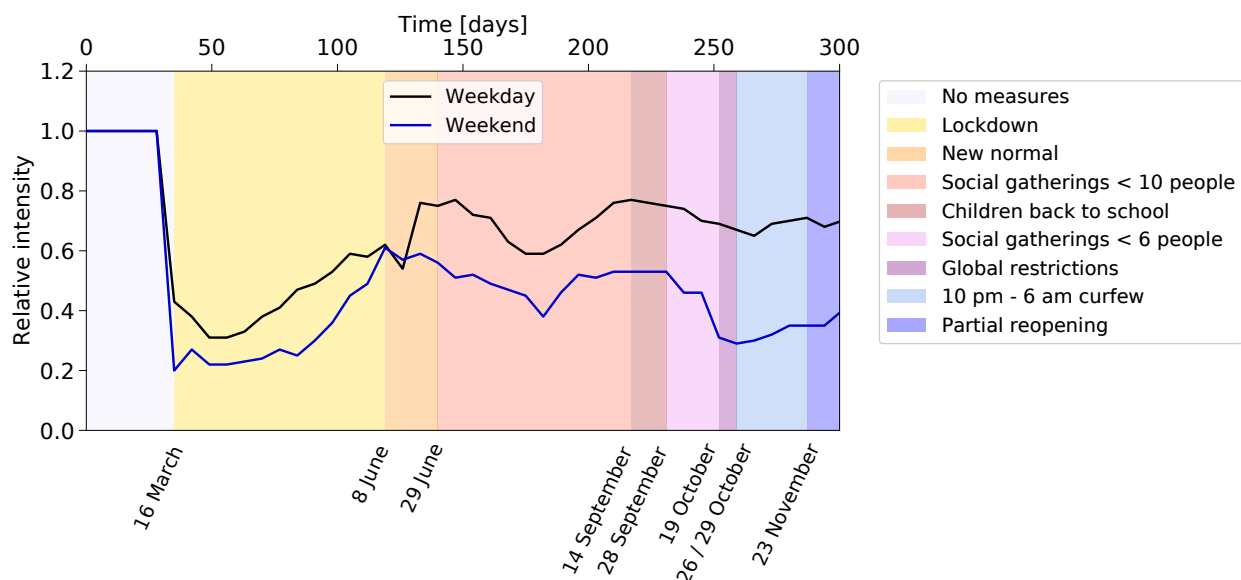


Figure 7: Relative intensity of mobile phones traffic and measures affecting work and community activities.

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.

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

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

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

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

397 **Sensitivity to Parameters**

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

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

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

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

407 duces the number of people being diagnosed, and subsequently put in quarantine. As a con-
408 sequence a factor of 2.5 more people become infected and the number of people diagnosed
409 increases by 30%.

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

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

421 Next we investigate the sensitivity to the pattern of contacts.

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

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

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

433 iii We studied the impact of making the number of community contacts dependent on the den-
434 sity of the population. We tested a model where the proportionality factor varies from 0.6 to
435 2, between Berguedà and Barcelonès, respectively, the lowest and highest density regions.
436 The average number of community contacts increases from 4.8 to 7.6. During the first wave,
437 the total number of diagnosed increases by 30%. The overall effect is less noticeable during
438 the second wave, but the number of diagnoses in the Barcelonès starts dominating, unlike
439 what is observed in Figure 5a. Such a strong density dependence is thus disfavoured.

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

447 (Figure 5a).

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

460 Finally, we studied variations in the description of mobility.

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

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

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

477 **Data availability**

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

489 **Code availability**

490 The code used in the simulation is available upon request.

491 **References**

- 492 1. P.G. Walker et al. The impact of COVID-19 and strategies for mitigation and suppression in
494 low- and middle-income countries. *Science* **369**, 6502, 413–422 (2020).
- 495 2. N.G. Davies et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and

- 496 demand for hospital services in the UK: a modelling study. *The Lancet* **7**, e375–e385 (2020).
- 497 3. The Computational Biology and Complex Systems group.
- 498 <https://biocomsc.upc.edu/en> (2021).
- 499 4. Agència de Qualitat i Avaluació Sanitàries de Catalunya (AQuAS).
- 500 <https://aquas.gencat.cat/ca/actualitat/ultimes-dades-coronavirus/> (2021).
- 501 5. Evolution of number of cases and R_t of SARS-CoV2, Agència de Qualitat i Avaluació Sanitàries
- 502 de Catalunya, (AQuAS).
- 503 [https://dlscitizens.blob.core.windows.net/rtreports/archived/20201231/CAT/](https://dlscitizens.blob.core.windows.net/rtreports/archived/20201231/CAT/InformeCasosRt_CAT_09.pdf)
- 504 [InformeCasosRt_CAT_09.pdf](https://dlscitizens.blob.core.windows.net/rtreports/archived/20201231/CAT/InformeCasosRt_CAT_09.pdf) (2020).
- 505 6. E. Estrada. Modeling the Present, Looking at the Future. *Physics Reports* **1869**, 1–51 (2020).
- 506 7. H. Khataee et al. Effects of social distancing on the spreading of COVID-19 inferred from
- 507 mobile phone data. *Scientific Reports* **11**, 1661 (2021).
- 508 8. E. Burns et al. The natural history of symptomatic COVID-19 during the first wave in Catalonia.
- 509 *Nature Communications* **12**, 777 (2021).
- 510 9. M. Castro et al., The turning point and end of an expanding epidemic cannot be precisely
- 511 forecast. *PNAS* **117** (42) 26190-26196 (2020).
- 512 10. C. O. Wilke and C. T. Bergstrom, Predicting an epidemic trajectory is difficult. *PNAS* **117** (46)
- 513 28549-28551 (2020).

- 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](https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176992&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).

- 533 19. Instituto Nacional de Estadística (INE). Studies on mobility based on mobile phone
534 https://www.ine.es/en/experimental/movilidad/experimental_em_en.htm (2020).
- 535 20. Instituto Nacional de Estadística (INE). Functional Urban Areas: Barcelona
536 <https://ine.es/index.htm> (2021).
- 537 21. J.M. Arauzo-Carod et al. Do local characteristics act in a similar way for the first two waves of
538 COVID-19? Analysis at intraurban level in Barcelona. *J Public Health (Oxford)* **43,3**, 455-461
539 (2021).
- 540 22. Q. Bi et al. Epidemiology and Transmission of COVID-19 in 391 Cases and 1286 of Their
541 Close Contacts in Shenzhen, China: A Retrospective Cohort Study. *The Lancet. Infectious
542 Diseases* **20,8**, 911–919 (2020).
- 543 23. X. He et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature
544 Medicine* **26**, 672–675 (2020).
- 545 24. L. Di Domenico et al. Impact of lockdown on COVID-19 epidemic in Île-de-France and pos-
546 sible exit strategies. *BMC Med* **18**, 240 (2020).
- 547 25. A. Aleta et al. Modelling the impact of testing, contact tracing and household quarantine on
548 second waves of COVID-19. *Nature Social Behaviour* **4**, 964-971 (2020).
- 549 26. R. Verity et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis.
550 *The Lancet. Infectious Diseases* **20,6**, 669-677 (2020).

- 551 27. Análisis de los casos de COVID-19 notificados a la RENAVE hasta el 10 de mayo en España.
552 Informe COVID-19 n° 33. 29 de mayo de 2020.
553 [https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/](https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx)
554 [EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx](https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx) (2020).
- 555 28. A. Soriano-Arandes et al. Household Severe Acute Respiratory Syndrome Coronavirus 2
556 Transmission and Children: A Network Prospective Study. *Clinical Infectious Diseases* **73,6**,
557 e1261–e1269 (2021).
- 558 29. T. Tolossa et al. Time to recovery from COVID-19 and its predictors among patients admit-
559 ted to treatment center of Wollega University Referral Hospital (WURH), Western Ethiopia:
560 Survival analysis of retrospective cohort study. *PLOS ONE* **16 (6)**, e0252389 (2021).
- 561 30. Centro de Ciencias Humanas y Sociales, CSIC. Envejecimiento en Red, datos de abril de
562 2019.
563 <http://envejecimiento.csic.es/documentos/documentos/enred-estadisticasresidencias2019.pdf>
564 (2019).
- 565 31. Instituto Nacional de Estadística (INE). Una estimación de la población que vive en residen-
566 cias de mayores.
567 [http://envejecimientoenred.es/una-estimacion-de-la-poblacion-que-vive-en-residencias-de-](http://envejecimientoenred.es/una-estimacion-de-la-poblacion-que-vive-en-residencias-de-mayores/)
568 [mayores/](http://envejecimientoenred.es/una-estimacion-de-la-poblacion-que-vive-en-residencias-de-mayores/) (2020).
- 569 32. Diputació de Barcelona. Informació Estadística Local.
570 <https://www.diba.cat/hg2/presentacioprov.asp?prid=954> (2020).

- 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/](https://educacio.gencat.cat/ca/departament/estadistiques/indicadors/sistema-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](https://ajuntament.barcelona.cat/estadistica/catala/Estadistiques_per_temes/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).

588 **Acknowledgments**

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

591 MB, XJ, LLM, MMar, PM, JP, IR acknowledge support from the grant 2020PANDE0180
592 of the programme PANDEMIES 2020, “Replegar-se per créixer: l’impacte de les pandèmies en
593 un món sense fronteres visibles” of the Agència de Gestió d’Ajuts Universitaris i de Recerca
594 of the Generalitat de Catalunya. LG thanks the funding from the European Union’s Horizon
595 2020 research and innovation programme under grant agreement ID 758145. AL acknowledges
596 the support from the Talent Research Program (Universitat Autònoma de Barcelona). PM has
597 received funding from the Spanish Ministry of Science and Innovation (PID2020-112965GB-
598 I00/AEI/ 10.13039/501100011033), from the Agency for Management of University and Research
599 Grants of the Government of Catalonia (project SGR 1069), and also received support from Ajunta-
600 ment de Barcelona. MMan acknowledges support from Marie Skłodowska-Curie grant agreement
601 ID 6655919. VV has been partially supported by the H2020 programme and by the Secretary
602 of Universities and Research of the Government of Catalonia through a Marie Skłodowska-Curie
603 COFUND fellowship – Beatriu de Pinós programme ID 801370.

604 **Author Contributions**

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

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

610 **Competing interests**

611 The authors declare no competing interests.

1 **Supplementary material.**

2 M. Bosman^{1*}, A. Esteve^{2,3}, L. Gabbanelli¹, X. Jordan⁴, A. López-Gay^{2,5}, M. Manera^{1,6}, M.

3 Martínez^{1,7}, P. Masjuan^{1,8}, Ll.M. Mir¹, J. Paradells^{4,9}, A. Pignatelli¹, I. Riu¹, V. Vitagliano^{1,10,11}

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

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

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

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

8 ⁵*Departament de Geografia, Universitat Autònoma de Barcelona, Bellaterra, Spain*

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

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

11 ⁸*Departament de Física, Universitat Autònoma de Barcelona, Bellaterra, Spain*

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

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

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

15 **Corresponding author: bosman@ifae.es*

16 **Parameters of the model**

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

19 are set according to external information. Out of the 64 parameters that control the simulation, only
20 8 are adjusted. Tables 1 to 7 describe the parameters, their settings and related references.

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

Category	Number	Parameters	Source	Adjusted
Disease				
Incubation time (Gamma distribution)	2	$\mu, \sigma = 4.58, 3.24$ days	Medical data ^{1, 2, 3}	No
Infectiousness time (Gamma distribution)	2	$\mu, \sigma = 2.5, 1.7$ 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 1 st 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 2 nd 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 2 nd wave

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 Oficial de la Generalitat de Catalunya ¹⁰	Shape of 1 st 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

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

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

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

21 **Statistics of the sample**

22 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
→ 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

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 municipality	BCN	< 20'	Home neighbourhood	190,316
		> 20'	Another BCN neigh.	254,219
	Counties	Any	Home region	448,229
Another municipality	BCN	Any	Counties	128,165
	Counties	< 45'	Home region	659,138
		> 45'	Outside home region	175,293
Several municipalities	BCN	Any	Counties	42,032
	Counties	Any	Home region	144,856
Abroad	Any	Any	Outside home region	7,278

23 Description of the province of Barcelona

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

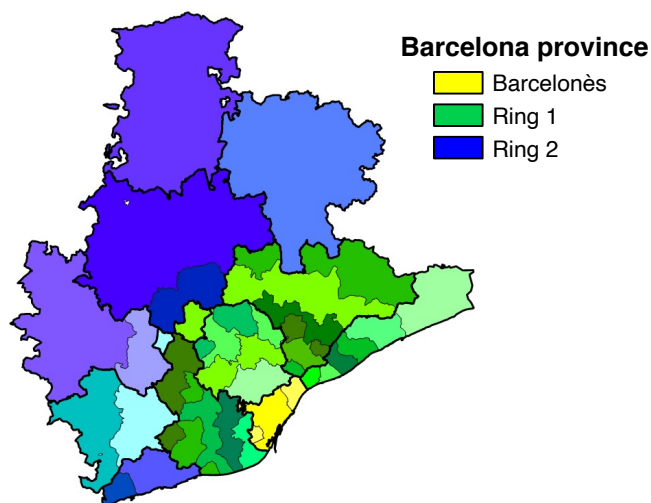


Figure 1: Map of the province of Barcelona. The regions used in the simulation regrouped in three concentric areas.

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

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,	33,018	el Parc i la Llacuna del Poblenou, la Vila Olímpica del	23,576
Provençals de Poblenou		Poblenou	
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	41,664
		la Guatlla	
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 Família (Nord)	22,534
el Baix Guinardó, Can Baró	36,252	la Sagrada Família (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	29,230
		d'Hebron, la Clota	
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		

34 **Age distribution of various sectors of the population**

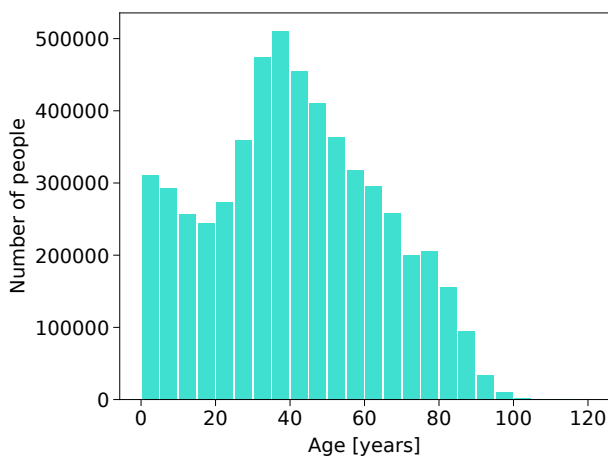


Figure 2: Population age. The age of the population of the province of Barcelona (5.5 million inhabitants) is shown.

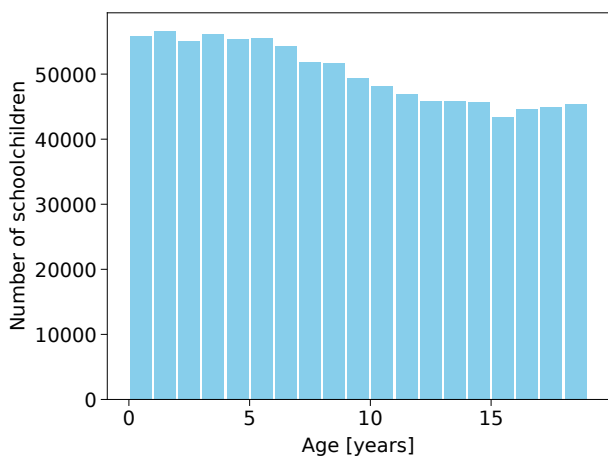


Figure 3: Schoolchildren age. The age of the 0 to 18-year-old children who attend school is shown.

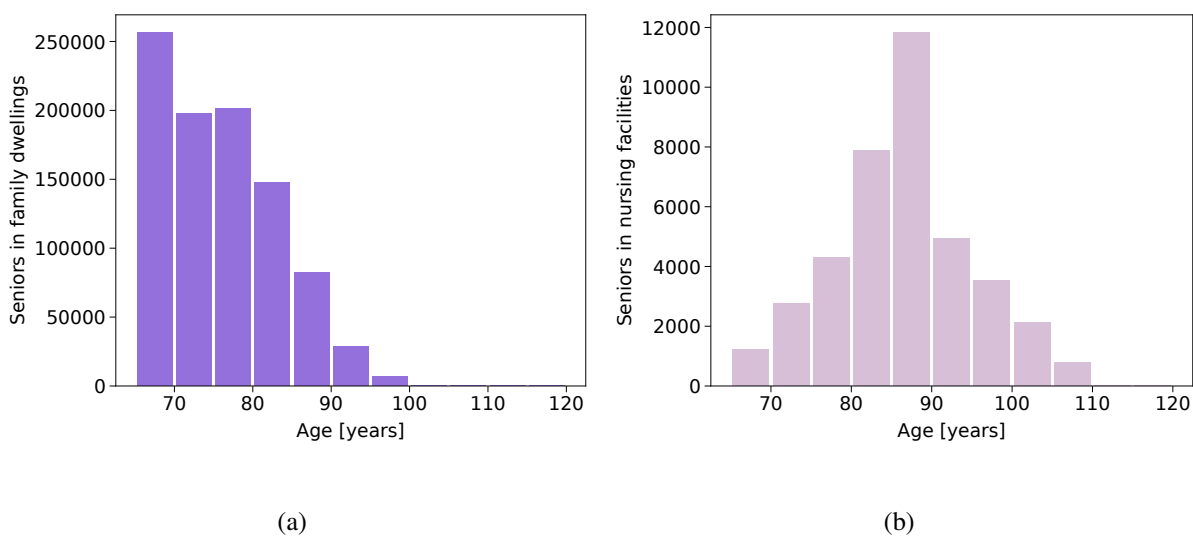


Figure 4: Seniors age. The age of people over 64 years of age is shown separately for **a** those who live in family dwellings, and **b** those who live in nursing facilities.

35 Time profile of incubation time and viral shedding

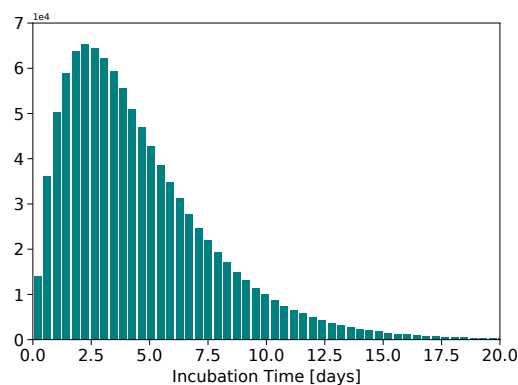


Figure 5: Incubation time. The time of incubation of the virus (in days) follows a Gamma distribution of mean 4.58 days and standard deviation 3.24 days.

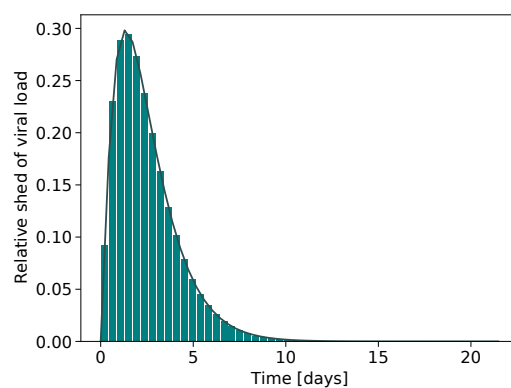


Figure 6: Relative profile of viral shedding. The distribution of the relative profile of the viral shedding follows a Gamma distribution of mean 2.5 days and standard deviation 1.7 days. In the simulation the distribution is truncated at 14 days.

- 36 1. Q. Bi et al. Epidemiology and Transmission of COVID-19 in 391 Cases and 1286 of Their Close
38 Contacts in Shenzhen, China: A Retrospective Cohort Study. *The Lancet. Infectious Diseases*
39 **20.8**, 911–919 (2020).
- 40 2. X. He et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature*
41 *Medicine* **26**, 672–675 (2020).
- 42 3. L. Di Domenico et al. Impact of lockdown on COVID-19 epidemic in Île-de-France and possi-
43 ble exit strategies *BMC Med* **18**, 240 (2020).
- 44 4. Análisis de los casos de COVID-19 notificados a la RENAVE hasta el 10 de mayo en España.
45 Informe COVID-19 n° 33. 29 de mayo de 2020.
46 [https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/](https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx)
47 [EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx](https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx) (2020).
- 48 5. T. Tolossa et al. Time to recovery from COVID-19 and its predictors among patients admitted to
49 treatment center of Wollega University Referral Hospital (WURH), Western Ethiopia: Survival
50 analysis of retrospective cohort study. *PLOS ONE* **16** (6), e0252389 (2021).
- 51 6. Departament d’Esenyament - Generalitat de Catalunya. Ràtios d’alumnes per estudi i unitat o
52 grup.
53 [https://educacio.gencat.cat/ca/departament/estadistiques/indicadors/sistema-](https://educacio.gencat.cat/ca/departament/estadistiques/indicadors/sistema-educatiu/escolaritzacio/ratios/)
54 [educatiu/escolaritzacio/ratios/](https://educacio.gencat.cat/ca/departament/estadistiques/indicadors/sistema-educatiu/escolaritzacio/ratios/) (2020).

- 55 7. K. Prem et al. Projecting contact matrices in 177 geographical regions: an update and compari-
56 son with empirical data for the COVID-19 era.
57 <https://doi.org/10.1101/2020.07.22.20159772> (2020).
- 58 8. Departament d'Estadística i Difusió de Dades - Ajuntament de Barcelona. Demanda hotelera a
59 Barcelona. Serie històrica 2005-2021.
60 [https://ajuntament.barcelona.cat/estadistica/catala/Estadistiques_per_temes/
61 Turisme_i_promocio_economica/Turisme/Oferta_demanda_hotelera/evo/th07.htm](https://ajuntament.barcelona.cat/estadistica/catala/Estadistiques_per_temes/Turisme_i_promocio_economica/Turisme/Oferta_demanda_hotelera/evo/th07.htm) (2021).
- 62 9. Enquesta de Mobilitat en Dia Feiner (EMEF) - 2019.
63 <https://www.atm.cat/web/es/observatori/encuestas-de-movilidad.php> (2019).
- 64 10. Diari Oficial de la Generalitat de Catalunya .
65 <https://dogc.gencat.cat/ca/inici/> (2020).
- 66 11. Y. Cheng et al. Face masks effectively limit the probability of SARS-CoV-2 transmission *Sci-*
67 *ence* **372**, **6549** 1439-1443 (2021).
- 68 12. Y. Wang et al. How effective is a mask in preventing COVID-19 infection? *Medical devices &*
69 *sensors* **4** e10163 (2021).
- 70 13. Estudios de movilidad a partir de la telefonía móvil.
71 https://www.ine.es/experimental/movilidad/experimental_em.htm (2020).