Testing and isolation to prevent overloaded healthcare facilities and reduce death rates in the SARS-CoV-2 pandemic in Italy

Arnab Bandyopadhyay1,4✉, Marta Schips1,4✉, Tanmay Mitra1, Sahamoddin Khailaie1, Sebastian C. Binder1 & Michael Meyer-Hermann1,2,3✉

Abstract

Background During the first wave of COVID-19, hospital and intensive care unit beds got overwhelmed in Italy leading to an increased death burden. Based on data from Italian regions, we disentangled the impact of various factors contributing to the bottleneck situation of healthcare facilities, not well addressed in classical SEIR-like models. A particular emphasis was set on the undetected fraction (dark figure), on the dynamically changing hospital capacity, and on different testing, contact tracing, quarantine strategies.

Methods We first estimated the dark figure for different Italian regions. Using parameter estimates from literature and, alternatively, with parameters derived from a fit to the initial phase of COVID-19 spread, the model was optimized to fit data (infected, hospitalized, ICU, dead) published by the Italian Civil Protection.

Results We show that testing influenced the infection dynamics by isolation of newly detected cases and subsequent interruption of infection chains. The time-varying reproduction number \( R_t \) in high testing regions decreased to <1 earlier compared to the low testing regions. While an early test and isolate (TI) scenario resulted in up to ~31% peak reduction of hospital occupancy, the late TI scenario resulted in an overwhelmed healthcare system.

Conclusions An early TI strategy would have decreased the overall hospital usage drastically and, hence, death toll (~34% reduction in Lombardia) and could have mitigated the lack of healthcare facilities in the course of the pandemic, but it would not have kept the hospitalization amount within the pre-pandemic hospital limit.

Plain language summary

Italy was heavily affected early in the COVID-19 pandemic, with healthcare facilities becoming overloaded. We use mathematical models to study COVID-19 transmission, and factors contributing to this, on a regional basis in Italy. We show that testing and lockdowns were effective in controlling disease spread. We use regional pre- and post-pandemic hospital/ICU bed occupancy to quantify the impact of the overwhelmed healthcare system upon the number of deaths. We find that increased isolation of cases could have reduced the effect of limited healthcare facilities but would not have kept hospitalizations within the pre-pandemic limit, and an improvement of hospital facilities would still have been required. We show that contact tracing and quarantine without testing could also be efficient strategies when test capacities are limited. Our findings help us to understand how to manage COVID-19 or other disease outbreaks in future.

1 Department of Systems Immunology and Braunschweig Integrated Centre of Systems Biology (BRICS), Helmholtz Centre for Infection Research, Braunschweig, Germany. 2 Institute for Biochemistry, Biotechnology and Bioinformatics, Technische Universität Braunschweig, Braunschweig, Germany. 3 Cluster of Excellence RESIST (EXC 2155), Hannover Medical School, Hannover, Germany. 4 These authors contributed equally: Arnab Bandyopadhyay, Marta Schips. ✉email: arnab.bandyopadhyay@theoretical-biology.de; marta.schips@theoretical-biology.de; mmh@theoretical-biology.de
The COVID-19 outbreak originated in Wuhan, People’s Republic of China, in early December 2019. Within two months, it erupted and unfolded with tremendous speed in Italy, which became the European epicenter of disease spreading, forcing the government to impose a lockdown on March 9th, 2020. On March 19th, 3405 people had already died in Italy, thereby surpassing China, while 41035 people were diagnosed as COVID-19 positive. This induced Italy to shut down all non-essential businesses on March 21th. Despite the strict measures applied, in Lombardia alone a total of 28545 symptomatic people were infected by April 8th, accounting for 12976 hospital admissions, followed by Emilia Romagna (4130), Piemonte (3196), and Veneto (1839). These large numbers led to the complete collapse of the healthcare system within a few weeks of the first detection of COVID-19 cases, most notably in Lombardia where even funeral homes had been overwhelmed and were incapable of responding in a reasonable time. Even though the state expanded the hospital and intensive care unit (ICU) capacities, it could not prevent the bottleneck situation of the healthcare system and presumably caused a large number of deaths for a prolonged period.

Many factors aggravated the COVID-19 situation in Italy, among which the distinct demographic structure of Italy with nearly 23% of the population of age 65 years or older, larger household size, and the prevalence of three-generation households compared to Germany as well as limited hospital and ICU capacities. At the beginning of the pandemic, Italy focused on testing symptomatic patients only, which resulted in a large proportion of positive tests and high case fatality rates (CFR) compared to other countries. A large proportion of cases remained undetected, which became a major driver of new infections. A different study estimates that in Italy the actual number of total infections was around 30-fold higher than reported, while for Germany it was less than ten-fold (data up to March 17th, 2020)

Compartmental models have been widely used to describe the dynamics of epidemics, for example, SIR models that consider three compartments, namely susceptible, infected, and recovered, or more complex SEIR models that take susceptible, exposed, infectious, and recovered compartments into account. Typically, these models either exclude the undetected index cases or ignore their dynamic nature, and structurally these models are not developed to address the load on the healthcare system. Besides these epidemic models, simple algorithms exist in the literature for estimating the time-varying reproduction number and have been widely used in the context of many infectious diseases (e.g., measles, H1N1 swine flu, polio, etc.). Several studies estimated the undetected case number in Italy, but its dynamics in the context of different testing strategies and implications on the healthcare system were not considered. Even though the general compartmental SIR and SEIR type models are useful in inferring epidemic spread and public health interventions, we needed to introduce additional compartments to investigate how the pandemic is shaped by several influential factors (e.g., dark numbers, regional testing strategies, hospital beds); for instance, we included a specific compartment for infected undetected cases ($I_D$) to analyse the impact of the region-wide undetected cases upon the evolution of the $R_0$. Similarly, hospital ($H_U$ and $H_R$) and ICU ($U_U$ and $U_R$) compartments were introduced to monitor the load on the healthcare system.

Additionally, the absence of reliable symptom onset data and heterogeneity in the actual infectious period among the asymptomatic, pre-symptomatic, and symptomatic individuals require a more complex model that not only can accurately portray the dynamics of COVID-19 spread but also can disentangle the impact of intertwined factors like the variation of undetected components, limited and changing hospital and ICU availability.

Existing modeling studies that analyze the COVID-19 situation in Italy, or other regions, in general, did not address some fundamental aspects of the ongoing pandemic like temporal dynamics of undetected infections, the benefits of a high testing and isolation strategy, or the impact of a limited and dynamically changing healthcare capacity on the lives lost. In this study, we address the bottleneck situation of the healthcare facility, the benefits of extending hospital infrastructure, and the impact of an early testing and isolation strategy on the healthcare system with a COVID-19-specific mathematical model. To evaluate the COVID-19 situation in Italy in a realistic framework, we first estimated the undetected fraction (dark figure) of infections across different regions of Italy. We used this information to determine the parameters of the model and showed that our model is structurally identifiable. We studied the influence of the dark figure and implemented NPIs on the time-dependent reproduction number, $R_0$. With data about regional hospital and ICU bed capacities, we estimated that an extra 25% of people died in Lombardia due to the overwhelmed healthcare system. We investigated the impact of early testing strategy and, alternatively, of contact tracing combined with quarantine (−10 fold more isolation of infected) policy in the setting of elevated hospital capacity as it currently stands. This strategy would have reduced the death toll by 20% to 50%.

Methods

SECRID-model. To understand the impact of potential aggravating factors, namely infections from undetected index cases, early vs late testing strategy, and limited healthcare facilities on disease progression, we developed a COVID-19-specific SECRID-model parameterized for Italy. The SECRID-model distinguishes healthy individuals without immune memory of COVID-19 (susceptible, $S$), infected individuals without symptoms but not yet infectious (exposed, $E$), and infected individuals without symptoms who are infectious (carrier, $C_I$, $C_R$). The carriers are distinguished into a fraction $\alpha$ of asymptomatic ($C_\alpha$) and $(1-\alpha)$ of pre-symptomatic infected ($C_\beta$). The latter are categorized into a fraction $\mu$ of detected symptomatic ($I_1$ and $I_2$) and $(1-\mu)$ of undetected mild-symptomatic ($I_3$). Out of the $C_I$, a fraction $\rho$ gets hospitalized ($I_1$), and $(1-\rho)$ become symptomatic but recover without hospitalization ($I_2$). Further, compartments for hospitalization ($H$) and intensive care units ($U$) were introduced to monitor the load on the healthcare system. A fraction $\vartheta$ of $H$ requires treatment in ICU ($H_1$) while a fraction $(1-\vartheta)$ recovers from hospital without ICU treatment ($H_2$). $\delta$ and $(1-\delta)$ represent the fraction of patients in ICU who subsequently die ($U_1$) or recover ($U_2$). The compartment ($R$) consists of patients recovered from different infection states. The Reference Model (Fig. 1; equations are in the Supplementary Methods 1.3) was solved with parameters in Table 1.
**Initial condition.** Italian regions started documenting epidemiological data at different dates, at the earliest February 24th, 2020. As we considered a fixed incubation period of 5.2 days in our model, we assumed that at minimum, the first entry in the dataset (number of total cases) was the exposed number 5.2 days earlier. In addition to the documented infection, we calculated undocumented cases based on the estimated region-wise dark number by a Bayesian MCMC framework (Supplementary Methods 1.2). The sum of documented and undocumented cases was the initial exposed population. We used regional population as the initial susceptible population. For all other compartments, we started the simulation with zero. The simulation began from \(-5.2\) days with the aforementioned initial conditions, and undetected cases were split between asymptomatic and symptomatic undetected cases according to the parameters used for those compartments.

**Parameterization.** We distinguished physiological and behavioral model parameters. Physiological parameters depend on the nature of the virus \(R_x, x = 2, \ldots, 9\) and remain unchanged throughout the analysis of the pandemic. We first determined the range of physiological values for each parameter from the literature\(^{27,28}\) (Table 1) and then estimated parameters’ values from a fit to the exponential growth of case numbers (infection, hospitalized, ICUs, and death cases) during the first two weeks of the pandemic. In total, 56 data points (14 daily data for those four observables) were used to estimate the physiological parameters. This initial phase was not yet affected by NPIs, public awareness, or an overwhelmed healthcare system and, thus, reflects viral properties. Some of the physiological parameters may be internally linked. For instance, hospitalization and ICU cases increase with infection cases, and disentangling those internal relations is difficult with limited data availability. It is likely that the best combination of parameters contains those internal relations. We assumed that the virus variant remained the same during the investigation period, therefore we kept physiological parameters constant throughout. However, environmental factors, testing policy, interventions, public behavior, self-isolation, hospitalization, etc. might have distorted such
This moving-window technique with the size of a calendar week region in Italy in moving time windows of 7 days to account for optimizer. This procedure was repeated separately for each randomly sampled parameters within a range specified and identiﬁed parameter sets for the statistical analysis. We generated dynamics for 100 periods of asymptomatic cases without requiring hospitalization (including the undetected symptomatic people (i_s)).

Table 1 Parameter ranges used in the Reference Model: determination of the boundaries for literature-based parameter set was based on the interpretation of the values given in the references27,28,32.

| Parameter | Comments/References | Description | Parameter ranges from literature |
|-----------|---------------------|-------------|---------------------------------|
| R_1       | Time-dependent      | transmission probability of COVID-19 per each contact made with an infectious person (C0, C0/k, k, and \( \lambda \) in the model) | Min Max |
| R_2       | 54,64-66            | the inverse of R_2 represents latent, non-infectious period following the transmission of COVID-19, 1/ R_2 = 5.2 – 1/ R_0; median incubation period is 5.2 days | \( \frac{1}{\gamma} \) \( \frac{2}{\gamma} \) |
| R_3       | 54,64-66 \( \delta \) | the inverse of R_3 represents the pre-symptomatic infectious period | \( \frac{1}{\gamma} \) \( \frac{2}{\gamma} \) |
| R_4       | 67-69               | the inverse of R_4 represents the infectious period for the mild symptomatic cases without requiring hospitalization (including the undetected symptomatic people (i_s)) | \( \frac{1}{\gamma} \) \( \frac{2}{\gamma} \) |
| R_5       | 27,28,70            | the inverse of R_5 represents the duration for which the hospitalized cases stay in general hospital care before discharge without requiring further intensive care | \( \frac{1}{\gamma} \) \( \frac{2}{\gamma} \) |
| R_6       | 28,71               | the inverse of R_6 represents the duration a patient stays at home before hospitalization | \( \frac{1}{7} \) \( \frac{2}{7} \) |
| R_7       | 28,70,71            | the inverse of R_7 represents the time spent in general hospital care before admission to ICU | \( \frac{1}{7} \) \( \frac{2}{7} \) |
| R_8       | 27,72               | the inverse of R_8 represents the time spent in ICU before recovery | \( \frac{1}{7} \) \( \frac{2}{7} \) |
| R_9       | 27,72               | the inverse of R_9 represents the time for which the asymptomatic cases remained infectious following their latent non-infectious period | \( \frac{1}{7} \) \( \frac{2}{7} \) |
| R_10      | Time-dependent28,71,73 | the inverse of R_10 represents the time spent in ICU before dying | \( \frac{1}{7} \) \( \frac{2}{7} \) |
| \( \alpha \) | fixed,30-38         | undocumented asymptomatic fraction | 0.4 0.4 |
| \( \beta \) | Assumed             | the risk of infection from the registered and quarantined (\( i_{h}+i_{0} \)) patients | 0.05 0.25 |
| \( \rho \) | Time-dependent       | the fraction of documented infections that require hospitalization | 0.01 0.9 |
| \( \theta \) | Time-dependent74-76 | the fraction of hospitalized patients that require further intensive treatment | 0.01 0.7 |
| \( \delta \) | Time-dependent74-76 | the fraction of ICU patients that have fatal outcome | 0.3 0.9 |
| \( \mu \)  | Bayesian framework   | this fraction represent the total undocumented infection including the asymptomatic cases, estimated through MLE method of the Bayesian framework | \( \frac{1}{\gamma} \) \( \frac{2}{\gamma} \) |

relations (e.g., the rate of increase in hospital and ICU cases as infection increases differ in different phases of the pandemic) and substantially altered the disease dynamics by impacting the transmission probability, dark number, hospitalizations, and death rate. These contingent factors affect the behavioral parameters \( (\rho, \theta, \delta, R_1, R_{10}) \). We estimated the behavioral model parameters by minimizing the sum of squared differences between the observed data (active infections, hospitalized, ICU patients, and death numbers (Italy Data on Coronavirus 202029)) and model simulations using Matlab’s nonlinear least-squares optimizer. This procedure was repeated separately for each region in Italy in moving time windows of 7 days to account for local specificities and temporal changes in disease transmission. This moving-window technique with the size of a calendar week reduces periodic fluctuations that are an artifact of the unequal distribution of tests among the weekdays.

Perturbation and parameter identifiability. To generate the standard deviation for \( R_1 \), we perturbed the behavioral parameters \( (\rho, \theta, \delta, R_1, R_{10}) \) 10% of their optimized value and sampled uniformly within this range such that the total parameter variation, \( \kappa \), defined as \( \log(\kappa) = \sum_{i=1}^{L} |\log k_i| \) [30], remains within 10% of its reference value. \( k_1, k_2 \) and \( L \) represent the parameters of the altered system, the reference system and the total number of parameters, respectively. We generated dynamics for 100 perturbed parameter sets for the statistical analysis.

We addressed parameter identifiability in two ways: Structural identifiability based on synthetic outbreak data and practical identifiability based on real data. In the first method, we randomly sampled parameters within a range specified in Table 1 and by using random initial conditions of model state variables. Then we used the resulting dynamics of the state variables as model observables and checked for a unique solution in the parameter space. We repeated this procedure 100 times (Supplementary Fig. 1 for a typical result). In the second method, we considered nationwide Italy data for the period February 24th to May 22nd, 2020, and fixed the physiological parameters as described in the Parameterization section. As we are estimating behavioral parameters \( (\rho, \theta, \delta, R_1, R_{10}) \) by considering a moving time windows of 7 days, we checked practical identifiability of these parameters in each time window. We found that the parameters are identifiable in more than 75% of the cases (Supplementary Fig. 2 for a typical results when all parameters are identifiable; and the Supplementary Methods 1.1 for more details).

Basic reproduction number. The basic reproduction number \( R_0 \) is defined as the expected number of secondary infections produced by a single infection in a population where everyone (assuming no immune memory) is susceptible[31] and reflects the transmission potential of a disease. For COVID-19, the dynamics of the pandemic was influenced by several factors, like, the self-awareness in the community, interventions and policies implemented by the authorities and immunisation of the population. Therefore, the time-dependent reproduction number \( R(t) \) that describes the expected number of secondary cases per infected person at a given time of the epidemic, is a more practically useful quantity to understand the impact of interventions, behavioral changes, seasonal effects, etc. on the disease dynamics[20,32].
In multi-compartmental epidemic models, $R_0$ can be derived with the next generation matrix method\textsuperscript{33–35}, where the Jacobian matrix consists of two factors, rate of appearance of new infections into the infection compartment ($F$) and transfer of infected into other compartments ($V$). The elements $G_{ij}$ of $G = F V^{-1}$ represent the expected number of secondary infections in compartment $i$ caused by a single infected individual of compartment $j$. The reproduction number $R_0$ is given by the dominant eigenvalue of $G$ (the derivation of $R_0$ is provided in the Supplementary Methods 1.4):

$$R_0 = \frac{S_0}{N_0} \left[ \frac{1 - \alpha - \beta \mu/C_0}{R_1} + \beta \mu + \frac{\alpha}{R_6} + \mu(1 - \rho) \frac{1 - \alpha}{R_1} + (1 - \mu) \frac{1 - \alpha}{R_1} \right].$$

(1)

where $N_0$ is the total population and $S_0$ is the susceptible population, both at the start of the pandemic (parameters are listed in Table 1). $R_0$ was calculated using the parameter set estimated by the initial fit that considers only the first two weeks of data points as described in the Parameterization section. An ensemble of parameter sets (as described in the Perturbation and parameter identifiability section) was used to calculate the standard deviation in the $R_0$.

In order to understand the impact of awareness in the population, NPIs and policies implemented by the authority upon the development of the time-varying reproduction number ($R(t)$)\textsuperscript{20}, we fitted the model parameters to data in shifting time windows of one week. This approach has two advantages: first, the reproduction number $R(t)$ is determined as a time-dependent variable and thus reflects the impact of NPIs on the infection dynamics; second, the moving-window dampens sudden jumps in the data because of reporting delays. In each time window, a best fit of the model parameters was found based on the cost function value (squared difference between data and simulation). In the next window, the fitting was repeated with initial conditions given by the model state in the previous time window. As described in the Perturbation and parameter identifiability section, an ensemble of perturbed parameter sets was used to calculate the standard deviation in the $R_0$.

$R(t)$ in time window $k$ reads:\textsuperscript{32}

$$R(t_k) = R_1(t_k) \frac{S(t_k)}{N(t_k)} \left[ \frac{1 - \alpha}{R_1} \right] + \frac{\beta \mu}{R_6} \left[ \frac{1 - \alpha}{R_1} \right] + \frac{\alpha}{R_6} \left[ \frac{1 - \alpha}{R_1} \right] + (1 - \mu) \left[ \frac{1 - \alpha}{R_1} \right] + (1 - \mu) \left[ \frac{1 - \alpha}{R_1} \right],$$

(2)

where $\rho(t_k)$ denotes the hospitalized fraction of identified symptomatic cases in the $k$th time window. Data analysis of the clinical state of all infected cases (up to June 22\textsuperscript{nd}) by the Istituto Superiore di Sanità (ISS) showed ~30% asymptomatic cases, with increasing tendency\textsuperscript{36–38}. In a study performed in Vo’ Euganeo, Veneto, the percentage of asymptomatic cases was found to be in the range of 40%\textsuperscript{39}. We set the asymptomatic fraction to $\alpha = 0.4$. The fraction of undetected cases $\bar{\mu}$ (Estimation of undetected cases section in the Supplementary Methods 1.2 and Table 2) is by definition:

$$\bar{\mu} := \alpha + (1 - \mu)(1 - \alpha) \Rightarrow \mu = \frac{1 - \bar{\mu}}{1 - \alpha}$$

(3)

Asymptomatic Model. In the Asymptomatic Model, all symptomatic cases are detected, i.e., $\mu = 1$. We compared the results from this model with those from the Reference Model to understand, in an ideal situation, the implication of detecting all symptomatic cases for the pandemic development.

Testing Model. In order to understand the influence of extra testing on infection dynamics, we adopted a model where a fraction of the undetected infected cases ($I_\text{SD}$) is detected ($I_\text{TD}$) via testing and hence, contained. The newly detected infected ($I_\text{TD}$) contribute to new infections with a frequency reduced by a factor $\beta$ ($\beta < 1$) but the infectious period remains unaltered ($1/\mu$) (Testing Model in Figs. 1 and 2). Li et al.\textsuperscript{40} have demonstrated, in the context of COVID-19 transmission in China, that strict control measures (travel restrictions, enhanced testing, self-quarantine, contact precautions, etc.) helped in improving the fraction of all documented infections from 14% to 65%, ~4–5 fold. For Italy, we estimated that 90% of the infections remained undetected (Table 2). We assumed that enhanced testing reduced this dark figure by daily 2% until reaching 60% (assuming similar efficiency as in China, i.e., documented infections increasing 10% to 40%). We introduced a time-dependent fraction $\mu'$ of undetected infections, which, starting from $\mu$, was decreased daily by steps of 2% down to 60%. The asymptomatic fraction was fixed as in the Reference Model ($\alpha = 0.4$). The undetected portion of symptomatic is instead modified so that the fraction of undetected cases, $\mu(t)$, and the fraction of newly detected cases, $\mu'(t)$, satisfies:

$$\mu_1(t) + \mu_2(t) = 1 - \mu,$$

(4)

with

$$\mu_1(t) = \frac{\mu'(t) - \alpha}{1 - \alpha}, \mu_2(t) = \frac{\bar{\mu} - \mu'(t)}{1 - \alpha}.$$

(5)

The parameters obtained by fitting the data with the Reference Model were transferred into the Testing Model. This maintains the compartmental flow of the Reference Model and thus ensures that the result reflects the sole effect of isolating a fraction of
undetected infections. Correspondingly, $R(t_k)$ becomes:

$$R(t_k) = R_1(t_k) \frac{S(t_k)}{N(t_k)} \left[ 1 - \alpha \frac{\beta \mu(t_k)}{R_8} + \frac{1 - \alpha}{R_9} + \alpha + \beta \mu(1 - \rho(t_k)) \frac{1 - \alpha}{R_4} + \beta \mu_2(t_k) \frac{1 - \alpha}{R_4} + \mu_1(t_k) \frac{1 - \alpha}{R_4} \right].$$  

(6)

Capacity Model. To estimate the impact of capacity limitations of the healthcare system, we implemented time-varying capacity constraints on the hospital ($H_{lim}(t)$) and ICU ($U_{lim}(t)$) accessibility (Capacity Model in Figs. 1 and 2), using available data for the number of hospital and ICU beds in the different regions. Table 3 reports the pre-pandemic capacity and the increased capacity, specifically allocated to COVID-19 patients, together with the date of accomplished installation. Some regions doubled their capacity and, presumably, this extension of infrastructure has been implemented in a step-wise manner. We assumed a linear increase of the hospital ($H_{lim}(t)$) and ICU ($U_{lim}(t)$) capacity from three days before exhaustion until reaching the maximum capacity on the date of accomplished installation. This new capacity was available thereafter. The exhaustion date was determined from the data and refers to the day at which the number of hospitalized and ICU patients became larger than the initial capacity. Before the pandemic, 85% of the hospital beds and 50% of the ICU beds were occupied. In the Capacity Model, 15% and 50% of the pre-pandemic total capacity (Table 3) was considered as the baseline capacity of hospital and ICU beds, respectively.

Upon reaching the capacity limit, the influx should be stopped until a vacant bed is available. This could be achieved by introducing a Heaviside step function or any other piecewise method, but this type of function introduces discontinuities and makes solving the ODEs computationally demanding and error-prone. Here, we introduced two functions ($f_{Hlim}$ and $f_{Ulim}$) that behave like a step function but are continuous in nature. $f_{Hlim}$ and $f_{Ulim}$ return 1 as long as there are free hospital or ICU beds and 0 otherwise. In the Capacity Model, we introduced the compartment $I_D$, to which the flux from $I_2$ is directed when the hospital capacity is reached. Then, in a sharp transition, patient access to the hospital was blocked.
the hospital or ICU is reduced by the fractions $f_{\text{Hlim}}(t)$ and $f_{\text{Ulim}}(t)$, respectively:

$$f_{\text{Hlim}} = \frac{\exp(H_R + H_U - H_{\text{Hlim}})}{1 + \exp(H_R + H_U - H_{\text{Hlim}})}^{10},$$  
$$f_{\text{Ulim}} = \frac{\exp(U_R + U_D - U_{\text{Ulim}})}{1 + \exp(U_R + U_D - U_{\text{Ulim}})}^{10}.$$

Both factors increase fatal outcomes of infections when hospital and ICU capacities are reached. We assumed that inaccessibility of hospital or ICU leads to faster and more frequent death. In particular, when the ICU capacity is reached, people in the hospital compartment ($H$) die after $1/R_T$ days which is faster than via the hospital-ICU-dead route ($1/R_T + 1/R_{\text{H}}$). Similarly, when the hospital capacity is reached, people in the infected compartment ($I$) die after $1/R_D$ days, satisfying $1/R_D < 1/R_R + 1/R_T < 1/R_R + 1/R_D + 1/R_H$ (see Reference Model in the Supplementary Methods 1.3).

The MaxCap Model is defined by the same equations as the Capacity Model, but the parameters, $H_{\text{Hlim}}$ and $U_{\text{Ulim}}$ were set to the maximum hospital and ICU capacity, respectively, from the beginning of the simulations.

### Data and code

Italy COVID-19 data of infected cases, hospitalized and ICU patients, and death numbers were provided by the Protezione Civile Italiana. Demographic and mortality data used to estimate IFR, are available from the Italian Institute of Statistics (ISTAT) website. ISTAT collects mortality data from the Italian National register office for the resident population (ANPR). An automated method was implemented, and parameter estimation was carried out in Matlab 2019b with a combination of the D2Dynamics framework. The code is available at [https://github.com/arnabbandyopadhyay/COVID-19-in-Italy](https://github.com/arnabbandyopadhyay/COVID-19-in-Italy), and has been archived on Zenodo at ref. 46. For the Bayesian estimation of COVID-19 IFR of Italian regions, see Supplementary Methods 1.2.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Results

#### Region-wise infection fatality rate (IFR)

To estimate undetected infections amid the COVID-19 pandemic, we analyzed the mortality rate of previous years and the deaths in 2020. Demographic and death data of the Italian regions have been collected from the Italian Institute of Statistics (ISTAT). The observed mortality of 2020 was substantially higher than in previous years in those Italian regions where the pandemic started—e.g., Lombardia, Veneto, Piemonte (Supplementary Fig. 3). We estimated the total number of infections, including undetected cases, and the associated IFR (defined as the percentage of deaths among all infections, including the undiagnosed infections) for each region with a Bayesian framework by adapting a standard binomial model (Supplementary Methods 1.2, Supplementary Fig. 4). Though literature is available on IFR estimates in Italy, the impact of the undetected infections on the healthcare facilities and infection dynamics, considering different testing strategies adopted by the regions, were not sufficiently addressed.

The regional IFR, the estimated total infections, Infection rate (IR), and the undetected fractions were summarised in Table 2. In the northern regions where the outbreak occurred first—e.g., Emilia Romagna, Piemonte, and Veneto—the undetected infections were nearly 10-fold more than the reported cases, and in Lombardia, it was more than 21 fold. We observed substantial heterogeneity of the IFR across different age groups (Supplementary Fig. 5). For ages below 60, it was as low as 0.05%. The IFR was drastically higher in the 81+ age group (9.5% to 20%, Supplementary Fig. 5). Despite Italy having the highest COVID-19 deaths in Europe, the estimated infection rates (IR) were relatively low (highest in Lombardia ~13%, 12.16–22.19%, 95%...
During the early outbreak, different northern regions in Italy adopted different testing strategies, which heavily influenced the infection dynamics. Lombardia and Piemonte followed the World Health Organization (WHO) and central health authority recommendations by mainly testing symptomatic cases, while Veneto implemented a more extensive population testing. Lombardia (~10 million inhabitants) has suffered around 14,000 deaths by the end of April, which is more than half of all COVID-19 deaths in Italy. In comparison, Veneto, with a population of 5 million, has suffered around 1,400 deaths. This difference is reflected in the CFR and IR values (Table 2): respectively, 6.43% and 2.77% in Veneto, while 18.30% and 13.83% in Lombardia, despite the geographical proximity.

To understand whether implementing different testing strategies succeeded in keeping undetected and the overall infection amount under control, we investigated the association between the number of tests performed per infection and total infections per capita. By considering 16 regions and Italy, a significant negative correlation confirms that the infection dynamics can be controlled by aggressive testing, which is further supported by the infection epidemic curve in Fig. 4a. Data are considered up to April 15th. Light blue line: linear regression fit; gray shaded area: standard error; black dots: region-specific values; red dot: nationwide value; R: correlation coefficient; p: significance.

Influence of undetected cases on \( R_t \). To understand the influence of undetected cases and installed NPIs on infection dynamics across different regions in Italy, we used the COVID-19-specific Reference Model to explain the dynamics of infected, hospitalized, ICUs, and death numbers provided by the Protezione Civile Italiana\(^{29}\) (Fig. 4a, Supplementary Fig. 7). As described in the Parametrization section, based on the ensemble of parameters estimated by using the first two weeks of data, we calculated the basic reproduction number \( R_0 \) according to equation (1). Estimated behavioral parameters in every window (keeping the physiological parameters constant) were used to calculate the reproduction number \( R_t \) (Fig. 4b, Supplementary Fig. 8) according to equation (2). The sudden increase in reported cases resulted in an overshoot in the \( R_t \) curve at the beginning. Due to the nationwide NPIs, increased public awareness inducing self-isolation and social distancing, the reproduction number continuously decreased, approaching unity at the end of April. In the regions with many undetected infections, like Emilia Romagna, Lombardia and Piemonte, the reproduction number reached unity in the first week of May, while in Veneto and Toscana it reached unity in the middle of April and was substantially lower by the first week of May. As \( R_t \) functionally depends on many factors (see \( R_t \) formula in the Methods section), we opted for a sensitivity analysis to find the important factors that regulate \( R_t \). It revealed that \( R_1 \) is highly sensitive to the transmission probability \( R_t \) and the dark figure \( \bar{\mu} \) (Supplementary Fig. 9). The influence of installed NPIs, social distancing, awareness etc. are embedded within \( R_1 \) and therefore, their impact is reflected in the decreasing trend of \( R_t \) in all regions. The benefits of testing and the impact of undetected cases on the \( R_t \) evolution can as well be inferred by comparison of \( R_t \) in the
Reference and the Asymptomatic Model with less undetected cases ($\mu = \alpha$, $\mu_1 = 1$).

In the Asymptomatic Model, all symptomatic infections are detected and, hence, this reduces secondary infections ($\beta << 1$). Removal of the highly infectious compartment ($I_X$), causes a lower turnover from susceptible to exposed. This effect is apparent in the initial phase of the pandemic when $R_t$ in the Asymptomatic Model is much lower than in the Reference Model (iris-blue and red line, respectively, in Fig. 4b). In the long term, the $R_t$ curves from the two models converge, despite the fraction of undetected cases being constant throughout the simulations: $\mu$ in the Reference Model and $\alpha$ in the Asymptomatic Model. This is consistent with the effect of the restriction measures, limiting the spreading through undetected cases. The difference in the $R_t$ curves of these two models is larger in the pre-lockdown period. As soon as people’s behavior has started changing either by self-awareness or by imposed restrictions, the turnover from susceptible to exposed is reduced. This caused the adjustment of $R_t$ and the merging of $R_t$ in both models. Thus, the influence of undetected infections on $R_t$ wanes (Fig. 4b) because of two reasons, first, restricted contacts due to the nationwide lockdown, and second, the enhanced testing strategy adopted by the regions.

**Fig. 4** Reference Model and time-dependent reproduction number $R_t$ in different regions of Italy. **a** Active infections, hospitalized, Intensive Care Units (ICU) and death data (nationwide and region-wise) were fitted in a sliding one week time window. Parameter ranges listed in Table 1 were used and the behavioral parameters ($R_t$, $R_{10}$, $\rho$, $\theta$, $\delta$) were estimated in each time window (see Methods); dots: data; continuous lines and shaded area: respectively, mean and standard deviation of all dynamics generated by using 100 perturbed parameter sets (see Methods). **b** Dynamics of the time-dependent reproduction-number, $R_t$, resulting from the fit with the Reference (red) and the Asymptomatic Model (blue). Statistics of $R_t$ were obtained by fitting the data with 100 perturbed parameter sets; continuous lines and shaded area are, respectively, mean and standard deviation. Vertical lines correspond to the Lockdown implementation (dark red) and release (dark green). Black horizontal line represents $R_t = 1$. 
Increased Test and Isolate (TI) strategy to reduce hospitalizations. As the impact of undetected cases on \( R_0 \), faded with time, we wanted to quantify the benefit of an increased test and isolate (TI) strategy by implementing an early (from March 2nd, 2020, i.e., one week before lockdown) and late (from March 15th, 2020) testing strategy. Many countries have implemented a Test, Trace and Isolate (TTI) strategy that has a high detection rate, and several modeling studies indicate that a high proportion of cases would need to isolate to control the pandemic.\(^{50,51}\) In our case, an increased test and isolate strategy with a high success rate of detection and isolation (as mentioned in the Methods section) is a phenomenological implementation of a TI strategy. In the Testing Model, we assumed that a fraction \( (\mu(t)) \) of the symptomatic undetected cases, \( I_u \), was detected by tests \( (I_{XD}) \) and, hence, became less infectious \( (\beta < 1, \text{Methods, Fig. 1}) \). We maintained the compartmental flow from the Reference Model by using the same parameter set in order to determine the impact of isolating undetected infections by targeted testing and home quarantine of contact clusters around identified infections. This in silico experiment resulted in a substantial increase in the number of detected infections but reduced the number of required hospitalizations. The early TI scenario resulted in up to \( \sim 32\% \) peak reduction of hospital bed occupancy, which reduced death numbers by up to \( \sim 44\% \) (Fig. 5a, b) depending on the region. The late TI scenario resulted in a situation similar to reality, namely, an overwhelmed healthcare system with little decrease in peak hospital occupancy. However, late TI still reduced death numbers by up to \( \sim 24\% \). Although enhanced testing increased the total number of detected infections in comparison to the real number of detected cases, \( R_0 \), fell and reached unity three weeks earlier (Supplementary Fig. 10). In line with previous results,\(^{50} \), this result suggests that TI strategies are efficient in decreasing disease propagation. South Korea successfully mounted a targeted testing strategy to contain disease spreading without imposing strict measures, like lockdowns or immigration control.\(^{8} \)

Having established that increased detection and isolation lower hospitalization rates, we investigated whether the same relation can be inferred from the data directly. We calculated the correlation between the median hospital occupancy and the total number of tests performed per infected up to May 22nd, 2020 (Fig. 5c). Regions with low testing were associated with the healthcare system hitting its capacity limits, while regions with intense testing kept a functional healthcare system. The northern regions of Italy faced bottleneck situations in hospitals, which is well reflected by the analysis (Fig. 5c, red color). Especially in Lombardia, where only symptomatic patients were tested, the healthcare system was overwhelmed. In contrast, Veneto performed \( \sim 4 \) times more tests per infected than Lombardia, which reduced the number of infections and hospitalizations.

In summary, the hospitalization and testing data resolved per region suggest the benefit of intense testing strategies to mitigate the load on the healthcare system. The in silico experiments add evidence that this relationship is induced by the interruption of infection chains, in particular, by the detection of undetected cases. Thus, testing not only improves knowledge on the infection state but also directly impacts the dynamics of the pandemic.

**Capacity Model and excess death due to the shortage of hospital beds.** Besides NPIs and promoting social distancing, self-isolation etc., strengthening the healthcare system is also an inevitable part of the government response. According to the data published by the Italian Ministry of Health, Italy had 3.18 beds per 1000 people with an average occupancy of 75-90% before the pandemic.\(^{34} \) Between March 1st and March 11th, 2020, 9–11% of the infected people were admitted to ICU. Out of total \( \sim 5200 \) ICU beds (pre-pandemic) in Italy, 2500 were already occupied by March 20th. To cope with this critical situation, each region increased the hospital and ICU facilities (Table 3). Despite the considerable increase in hospital and ICU capacity, the unexpected huge wave of patients and the necessary time to adapt the facilities added to the difficulties of crisis management.

To investigate the impact of the limited healthcare capacities, we developed the Capacity and the MaxCap Models (Figs. 1 and 2). As of May 22nd, all regions had increased the number of available beds, and the epidemic curves were in a downward phase. Therefore, we investigated a possible impact on the pandemic of hospital overload in the previous months, considering data from February 24th to May 22nd.

In the Capacity Model, the pre-pandemic occupancy of hospital beds and ICUs was considered 85% and 50%\(^{11} \), and the baseline number of available hospital and ICU beds was set to 15% and 50% of the total capacity (Table 3), respectively. The capacity that was increased during the crises was described by a region-specific linear function with a daily increment so that the capacity reached its target value at the date indicated in Table 3. We assumed that the unavailability of hospital or ICU beds leads to faster and more frequent death (see Methods). Parameters were determined using the same protocol as for the Reference Model. The Capacity Model fit is shown in Fig. 6a and Supplementary Fig. 11.

In the MaxCap Model, the hospital and ICU capacities were fixed to their maximum from the very beginning. The compartmental flow in both models was kept identical by using the same parameter set as for the Capacity Model. We quantified the impact of the limited capacity on COVID-19-associated deaths by subtracting the death numbers in the MaxCap Models from those in the Capacity Model. This difference represents the number of people that would have benefited from a system with a substantially higher capacity at the beginning of the epidemic. The effect of limited healthcare facilities was dramatic in Lombardia with a \( \sim 26\% \) difference in the number of deaths, corresponding to \( \sim 4500 \) people (Fig. 6b).

A combined strategy. In reality, many regions ramped up facilities to test the immediate contacts of an identified infection, e.g., Veneto\(^{32} \) and also strengthen healthcare facilities to accommodate more patients. Previously we observed the benefits of an early TI strategy that reduces the load upon the healthcare system, thereby reducing death numbers (Testing Model, Fig. 5a, b). We also observed a reduction in fatal outcomes in the case of a functional healthcare system that is not overwhelmed (Capacity Model, Fig. 6a, b). Given these observations, we sought to investigate the combined effects of an improved healthcare facility with intense testing by considering the following four scenarios:

1. linear increase of hospital capacity combined with early TI;  
2. linear increase of hospital capacity combined with late TI;  
3. maximum hospital capacity from the beginning combined with early TI;  
4. maximum hospital capacity from the beginning combined with late TI.

Thereby, early/late was assumed one week before/after the lockdown.

To simulate scenarios 1 and 2, we transferred the Capacity Model parameters into the TestCap Model (Fig. 2) to keep the compartmental flow intact. To simulate scenarios 3 and 4, we evaluated the TestCap Model assuming the hospital and ICU capacity at their maximum levels from the beginning. In all
scenarios, improved TI was simulated as a step-wise reduction of undetected cases by 2% per day \( (\mu(t), \text{Testing Model in Methods}) \), which is equivalent to \(~10\)-fold increased detection.

Figure 7 reveals several interesting facets. The importance of testing in a realistic situation where the step-wise extension of the healthcare facility has been installed is depicted in Fig. 7a. It reveals that an early TI strategy brings down hospitalizations close to the pre-pandemic hospital capacity (horizontal black line). An exception is Lombardia, where adopting an early TI strategy led to \(~10\)% reduction of the hospitalization peak, while the late TI strategy did not decrease the peak size and led to a situation similar to what happened in reality. Though the peak size remained unchanged in Lombardia in the late TI scenario, simulations showed a reduction in the number of deaths of \(~17\)%.
reaching ~40% (Fig. 7b) when combined with an increased hospital capacity. However, adopting an early TI strategy is effective in reducing the death toll across all regions, ranging from ~34% to ~52%.

Figure 7B represents the importance of a TI strategy in the settings of a functional healthcare system. Previously, we estimated a 26% reduction in death numbers in Lombardia by strengthening the hospital infrastructure (Fig. 6b). Improved hospital capacity with early TI further reduced death numbers by a substantial amount, which ranged from ~35% to 52% across different regions, ~52% in Lombardia (Fig. 7b). Early TI with ~5-fold more testing would have reduced the death toll up to ~33% in Lombardia (Supplementary Fig. 12). In the late TI scenario, it would have decreased ~40% in Lombardia (Fig. 7b), and considering only ~5 times more testing it would have decreased deaths by ~28% (Supplementary Fig. 12).

In summary, this in silico experiment emphasizes the importance of an early TI strategy that could partially compensate for limited healthcare facilities during the early period (March to May 2020) of the pandemic. However, such an early TI strategy would not be sufficient to contain the hospitalizations within the pre-pandemic hospital limit (horizontal black line, Fig. 7). Therefore extending the hospital infrastructure was mandatory to prevent an overwhelmed healthcare system.
Discussion

Our retrospective study revealed many lessons that can be learned from the COVID-19 situation in Italy. Though containment measures are necessary to reduce the exponential growth of the pandemic and flatten the infection curve as early as possible, prolonged lockdown (mass quarantine) is not a sustainable solution to the pandemic given its socio-economic burden. The power of lockdown lies in restricting the contacts and reducing the spread mainly from the asymptomatic and pre-symptomatic carriers. This particularly applies to COVID-19 because a substantial part of secondary infections occurs prior to disease onset. Several studies indicate that the silent transmission from pre- and asymptomatic patients was responsible for the majority of new infections. Although a recent study has found that there is a lower risk of transmission from asymptomatic people, asymptomatic cases still present a public-health risk, as they are usually unaware of their infection and do not take any increased self-isolation measures. In Fig. 3 we showed that testing not only reduces the dark figure but also interrupts subsequent transmission chains from undetected cases and thereby directly influences the infection dynamics. This emphasizes the importance of a massive testing strategy to control the pandemic.

With the help of a mathematical analysis of the different regions of Italy, we have provided evidence for a general relation between intense testing and reduced burden on the healthcare system. To interrupt virus transmission chains, the Veneto
Region developed a comprehensive public health strategy focusing on case finding, contact tracing and quarantining close and occasional case contacts. Moreover, testing was extended to both symptomatic and non-symptomatic case contacts. Toscana followed a similar strategy as Veneto, whereas Lombardia tested only the symptomatic cases. Lombardia has twice the population of Veneto (10M vs 5M). Tests performed per capita in Veneto were almost twice as high as in Lombardia. The epidemiological outcomes of the testing strategy adopted by Veneto and Lombardia were different in terms of incidence number, evolution of the pandemic and bottleneck situation of the healthcare system. Veneto and Toscana flattened the infection curve one month earlier than Lombardia and Piemonte. In Fig. 3, Lombardia, Emilia, and Piemonte are in the low test, high infection regime, whereas Veneto is in the high test and low infection regime. Therefore, the pandemic situation in different Italian regions must be addressed in light of their testing strategy. Different testing strategies and their implications for the reproduction number have been studied. Several other studies emphasize the combination of contact tracing with testing and its implication on the incidence number. Our results suggest that the bottleneck in the healthcare system of northern Italian regions was a consequence of their TTI strategy. In particular, testing of symptomatic individuals alone appears inefficient. In the middle of the crisis, many healthcare workers were infected with COVID-19 while treating COVID-19 patients, and the voluntary participation of interns and retired personnel was required. Moreover, delays in the testing of healthcare personnel led to the spread of infection through healthcare workers. The shortage of healthcare workers together with limited hospital beds led to a bottleneck situation in the healthcare system. A study revealed that a weekly screening of the healthcare workers and other high-risk groups irrespective of their symptoms would reduce transmission by 23%. We quantified the impact of the overwhelmed healthcare system on the death toll and studied how such a bottleneck situation could be avoided. We showed that a substantial portion of the death toll, ~35% in Lombardia, could be prevented by testing, and this could have mitigated the shortage of healthcare facilities at the early stage of the pandemic, though it would not have contained the hospitalization within its pre-pandemic limit in all cases. Hence, given the capacities, the bottleneck of healthcare facilities could not be completely avoided by adopting massive testing alone.

A ~10-fold increased testing demands huge testing facilities and is not an economic strategy. Moreover, successful detection depends on the subjects getting tested and their previous travel and contact history. Alternatively, tracking and targeted testing with the home quarantine of possible cases substantially reduce infection transmission and could be adopted instead of rapid mass testing due to the higher chance of successful detection. However, manual contact tracing is infeasible at higher incidences, and contact tracing apps require ethical clearance in accessing the location data, transparency and protection of personal data. In this context, it has to be emphasized that, in the framework of our mathematical model, the impact of tests on the spreading of infections is based on the isolation of positively tested individuals regardless of their symptoms, in line with previous observations. This implies that contact tracing and quarantine without testing would have a similar effect and might be an efficient strategy when sufficient test capacities are not available. Thus, an effective contact tracing and quarantine mechanism, monitored through modern technologies, together with improved healthcare facilities could reduce mortality in the possible future waves or other pandemics.

Our modelling study has several limitations. During the first wave in Italy, there was a high degree of uncertainty regarding the fraction of pre-symptomatic and asymptomatic cases and their associated transmission. These fractions were also subject to the regional testing strategy, and their dynamic nature is observed in the weekly reports published by ISS. In our analysis, we set the asymptomatic fraction, $\alpha$, to the national average, 0.4. In our simulation, we considered a fixed incubation period. We performed a sensitivity analysis of $R_t$ upon the $R_e$ within the range provided in Table 1, and we found that $R_e$ is less sensitive towards the variations of $R_t$.

As mentioned in the Parameterization section, we determined the range of parameter values that constrained the fit of the physiological parameters by the information available in the literature on the virus characteristics. Some of the physiological parameters in the model may be internally linked. For example, hospitalization and ICU cases increase with infection cases. Here, we assumed that the nature of the virus did not change considerably during the investigation period, and, therefore, we kept the physiological parameters, and hence the internal relations, constant throughout the investigation period. The behavioral adjustments (through NPIs and awareness) caused the breakdown of such relationships and shaped the pandemic. As the behavioral parameters reflect the public behavior which evolved with time, they depend upon several factors and are region-specific. Therefore, it is difficult to infer a proper distribution of such parameters. Nevertheless, as a more precise parameter distribution would become available, sampling from a different distribution might improve the quality of the model.

For the present analysis, we did not use an age-stratified version of the Reference Model due to the lack of knowledge of age-stratified model parameters and incomplete age-stratified data during the first wave. However, the age-dependence is phenomenologically included in the model by using a time-dependent hospitalization rate, which reflects the demography of the infected people and other contingent factors that might alter the outcome of the pandemic. We also did not include co-morbidities or pre-existing medical conditions of a sub-population and did not explicitly consider the potential changes in viral transmissibility due to environmental factors, such as temperature and humidity. Further, cross-regional movements and the potential for imported or exported infections, which might hamper the testing strategy based on contact tracing, were not considered. Lastly, a delay in testing results might hamper outcomes in a contact tracing-based testing strategy. In our simulation, we did not explicitly consider delays in isolation and their impact on the daily cases. Instead of addressing the above limitations explicitly, we perturbed our behavioral parameter set up to 10% of its base value to ensure the robustness of our results in a plausible range of parameter values.

During the first wave of COVID-19, hospital and ICU beds got overwhelmed in Italy. There are potentially many factors, such as infections from undetected index cases, early vs late testing strategies, limited healthcare facilities, etc., that might have aggravated the COVID-19 situation in Italy. In this paper, we developed a COVID-19 specific infection epidemic model to address the bottleneck situation of the healthcare system that most of the northern regions of Italy, particularly Lombardia have faced during the first wave of the pandemic. As the testing was minimal at the beginning of the pandemic, a large portion of cases remained undetected, which was a major driver of new infections. We first estimated the dark figure for different regions of Italy through a Bayesian Markov Chain Monte Carlo (MCMC) framework. With an adaptive methodology, we estimated the model parameters by fitting the active cases, hospitalized, ICU, and death data published by the Civil Protection Department, Italy. We showed that testing directly influences the infection dynamics by interrupting transmission from undetected cases.
We showed that intense testing is associated with a reduced burden on the healthcare facility and, in reality, regional hospitals were less overwhelmed with more testing. By considering regional pre- and post-pandemic hospital and ICU beds we quantify the impact of the overwhelmed healthcare system on the death amount. This impact was highest in Lombardia, which affected ~4000 people. Implementing an early TTI strategy in Lombardia would have decreased overall hospital occupancy, which would have reduced the death toll by ~45%. However, such a strategy would not have kept the hospitalization amount within the pre-pandemic hospital limit. In this context, it is important to keep in mind that the effectiveness of such a strategy lies in the isolation of positively tested individuals regardless of their symptoms. Therefore contact tracing and quarantine without testing would have a similar effect and might be an efficient strategy when sufficient test capacities are not available.

Data availability
The dataset analysed during the current study are available in the github repository46. Source data for generating the figures are available inside the ‘PARALLEL/Reference_Asymptomatic’ and ‘PARALLEL/Capacity’ folders. Additional instructions to generate figures are given in the readme file in those folders.

Code availability
The code used for the analysis is available in the GitHub repository: https://github.com/arnabbandyopadhyay/COVID-19-in-Italy, and has been archived on Zenodo46.

Received: 2 September 2021; Accepted: 10 June 2022; Published online: 27 June 2022

References
1. worldometers-Coronavirus (2020). https://www.worldometers.info/coronavirus/. 2. Hollingsworth, T. D., Klinkenberg, D., Heesterbeek, H. & Anderson, R. M. Mitigation strategies for pandemic influenza A: balancing conflicting policy objectives. PLoS Comput. Biol. 7, e1001076 (2011). 3. Yang, P. et al. The effect of multiple interventions to balance healthcare demand for controlling COVID-19 outbreaks: a modelling study. medRxiv 2020.05.19.20107326 (2020). https://www.medrxiv.org/content/medrxiv/early/2020/05/26/2020.05.19.20107326.full.pdf. 4. Diehm, J. et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. Science 369 (2020). https://www.ncbi.nlm.nih.gov/pubmed/32414780. 5. Khailaie, S. et al. Development of the reproduction number from coronavirus sars-cov-2 case data in germany and implications for political measures. BMC Med. 19, 1–16 (2021). 6. Dietz, K. The estimation of the basic reproduction number for infectious diseases. Stat. Methods Med. Res. 2, 23–41 (1993). 7. Diekmann, O., Heesterbeek, J. & Metz, J. A. On the definition and the transmission of contact numbers in compartmental epidemic models. Math. Biosci. 180, 29–48 (2002). 8. ISTAT: Decesoi e cause di morte: cosa produce l’indice di popolazione. (2019). https://www.istat.it/it/dati/934(2020). 9. Barkai, N. & Leibler, S. Robustness in simple biochemical networks. Nature 387, 931–97 (1997). 10. van den Driessche, P. & Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. 180, 29–48 (2002). 11. 11. ISTAT: Sorveglianza integrata COVID-19 in Italia. (2020). https://www.epicentro.iss.it/coronavirus/bollettino/Infogra菲_c_22giugno%20ITA.pdf. 12. ISTAT: Sorveglianza integrata COVID-19: archivio (2020). https://www.epicentro.iss.it/coronavirus/sars-cov-2-sorveglianza-dati-archivio. 13. Lavezzo, E. et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vòo. Nature 584, 425–429 (2020). 14. Li, R. et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov-2). Science 368, 489–493 (2020). 15. Annaurario Statistico del Servizio Sanitario Nazionale. http://www.salute.gov.it/imgs/C_17_pubblicazioni_2879_allegato.pdf (2017). 16. ISTAT: Decessi e cause di morte: cosa produce l’Istat (2020). https://www.istat.it/it/archivio/240401 (Health, Demography, Society data) Last accessed: 31 July, 2020. 17. ISTAT: ISTAT-popolazione-e-famiglie. (2019). https://www.istat.it/it/popolazione-e-famiglie/dati (Mortality data) Last accessed: 31 July, 2020. 18. Analysis JHUM. https://coronavirus.jhu.edu/data/mortality (2020). 19. Bonmer, C. & Vollmer, S. Estimates of the severity of coronavirus disease 2019-ncov in the united states. University of Goettingen (2020). https://www.uni-goettingen.de/de/document/download/0af1dca623053908de337e1045c6612.pdf/COVID-19%20update.pdf. 20. Kermack, W. O. & Mc Kendrick, A. G. A contribution to the mathematical theory of epidemics. Proc. Royal Soc. Lond. Ser. A. Contain. Pap. Math. Phys. Eng. Sci. 115, 700–721 (1927). 21. Lin, Q. et al. A conceptual model for the coronavirus disease 2019 (Covid-19) outbreak in wuhan, china with individual reaction and governmental action. Int. J. Infect. Dis. 93, 211–216 (2020). 22. Anastasopoulou, C., Russo, L., Tsatsakis, A. & Siettos, C. Data-based analysis, modelling and forecasting of the covid-19 outbreak. PLoS One 15, e0230405 (2020). 23. Gatto, M. et al. Spread and dynamics of the covid-19 epidemic in italy: effects of emergency containment measures. Proc. Natl. Acad. Sci. 117, 10484–10491 (2020). 24. Cori, A., Ferguson, N. M., Fraser, C. & Cauchemez, S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am. J. Epidemiol. 178, 1565–1572 (2013). 25. Modi, C., Böhm, V., Ferraro, S., Stein, G. & Seljak, U. Estimating covid-19 mortality in italy early in the covid-19 pandemic. Nat. Commun. 12, 1–9 (2021). 26. Vollmer, M., Mishra, S., Juliette, H. et al. Report 20: Using mobility to estimate the transmission intensity of COVID-19 in Italy: a subnational analysis with future scenarios. Imperial College London. Imperial College London (2020). https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19-report-20-italy/. 27. Giordano, G. et al. Modelling the covid-19 epidemic and implementation of population-wide interventions in italy. Nat. Med. 26, 855–860 (2020). 28. IMHE COVID-19 Forecasting Team. Modeling COVID-19 scenarios for the United States. Nat. Med. 27, 94–105 (2021). 29. Davies, N. G. et al. Effects of non-pharmaceutical interventions on covid-19 cases, deaths, and demand for hospital services in the uk: a modelling study. Lancet Public Health 5, e375–e385 (2020). 30. Kucharski, A. J. et al. Early dynamics of transmission and control of covid-19: a mathematical modelling study. Lancet Infect. Dis. 20, 533–538 (2020). 31. RKI Steckbrief. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html (2020). 32. The epidemiologia per la sanità pubblica, I.d.s. characterisitcs of SARS-CoV-2 patients dying in Italy Report based on available data on June 4th, 2020. https://www.epicentro.iss.it/en/coronavirus/bollettino/Report-COVID-19_4_june_2020.pdf (2020). 33. COVID-19 Italia - Monitoraggio situazione. https://github.com/pcm-dpc/COVID-19 (2020). 34. Tai, B. & Leibler, S. Robustness in simple biochemical networks. Nature 387, 931–97 (1997). 35. Dietz, K. The estimation of the basic reproduction number for infectious diseases. Stat. Methods Med. Res. 2, 23–41 (1993). 36. Khailaie, S. et al. Development of the reproduction number from coronavirus sars-cov-2 case data in germany and implications for political measures. BMC Med. 19, 1–16 (2021). 37. Diekmann, O., Heesterbeek, J. A. & Metz, J. A. On the definition and the computation of the basic reproduction ratio $R_0$ in models for infectious diseases in heterogeneous populations. J. Math. Biol. 28, 363–382 (1990). 38. Diekmann, O., Heesterbeek, J. A. P. & Roberts, M. G. The construction of next-generation matrices for compartmental epidemic models. J. R. Soc. Interface 7, 873–885 (2010). 39. van den Driessche, P. & Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. 180, 29–48 (2002). 40. ISTAT: ISTAT-popolazione-e-famiglie. (2019). https://www.istat.it/it/dati/934(2020).
44. MATLAB: version: 9.7.0 (R2019b) (The MathWorks Inc., Natick, Massachusetts, 2019). https://www.mathworks.com/

45. Raue, A. et al. Data2Dynamics: a modeling environment tailored to parameter estimation in dynamical systems. Bioinformatics 31, 3558–60 (2015).

46. Bandypadhyay, A. et al. arnabbandypadhyay/covid-19-in-italy: Covid-19-in-italy (v1.1) (2022). https://zenodo.org/record/6497184.

47. Anderson, R. M. & May, R. M. Vaccination and herd immunity to infectious diseases. Nature 313, 333–339 (1985).

48. Randolph, H. E. & Barreiro, L. B. Herd immunity: understanding covid-19. Immunity 52, 737–741 (2020).

49. Alimohamadi, Y., Taghdir, M. & Sepandi, M. Estimate of the basic reproduction number for covid-19: a systematic review and meta-analysis. J. Prev. Med. Public Health 53, 151 (2020).

50. Kucharski, A. J. et al. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of sars-cov-2 in different settings: a mathematical modelling study. Lancet Infect. Dis. 20, 1151–1160 (2020).

51. Hellewell, J. et al. Feasibility of controlling covid-19 outbreaks by isolation of cases and contacts. Lancet Glob. Health 8, e488–e496 (2020).

52. Burn n. 37 del 20 marzo 2020, Materia: Sanità e igiene pubblica. https://bur.regione.veneto.it/Bur/Servizi/pubblica/DettagliDByr.aspx?id=417291 (2020).

53. International Labour Organization. Impact of lockdown measures on the informal economy A summary (2020). https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/travail/documents/briefingnote/wcms_743554.pdf.

54. Nishiura, H., Linton, N. M. & Akhmetzhanov, A. R. Serial interval of novel coronavirus (COVID-19) infections. Int. J. Infect. Dis. 93, 284–286 (2020).

55. Ferretti, L. et al. Quantifying sars-cov-2 transmission suggests epidemic control with digital contact tracing. Science 368 (2020).

56. Moghadas, S. M. et al. The implications of silent transmission for the control of covid-19 outbreaks. Proc. Natl. Acad. Sci. 117, 17513–17515 (2020).

57. Bi, Q. et al. Insights into household transmission of sars-cov-2 from a population-based serological survey. Nat. Commun. 12, 3643 (2021).

58. Byambasuren, O. et al. Estimating the impact of asymptomatic covid-19 and its potential for community transmission: systematic review and meta-analysis. Offic. J. Assoc. Med. Microbiol. Infect. Dis. Can. 5, 223–234 (2020).

59. Russo, F. et al. Epidemiology and public health response in early phase of covid-19 pandemic, veneto region, italy, 21 february to 2 april 2020. Eurosurveillance 25, 2000548 (2020).

60. Romagnani, P. et al. The covid-19 infection: lessons from the italian experience. J. Public Health Policy 41, 238–244 (2020).

61. Grassey, N. C. et al. Comparison of molecular testing strategies for covid-19 control: a mathematical modelling study. Lancet Infect. Dis. 20, 1381–1389 (2020).

62. Quilty, B. J. et al. Quarantine and testing strategies in contact tracing for sars-cov-2: a modelling study. Lancet Public Health 6, e175–e183 (2021).

63. Kretzschmar, M. et al. Impact of delays on effectiveness of contact tracing in different settings: a mathematical model and protocols for parameter estimation of SECIRD models. T.M. developed other versions of the SECIRD model. T.M. derived the literature-based parameter sets and analytical expressions of the reproduction numbers of the models. T.M. and S.K. provided insights through discussion. A.B. wrote the paper. A.B., M.M.H., M.S., and S.B. revised the paper. M.M.H. supervised the study.

64. Li, Q. et al. Early transmission dynamics in Wuhan, China, of Novel Coronavirus-infected pneumonia. N. Engl. J. Med. 382, 1199–1207 (2020).

65. Wang, D. et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in wuhan, china. JAMA 323, 1061–1069 (2020).

66. Zhao, S. et al. Estimating the serial interval of the novel coronavirus disease (covid-19): A statistical analysis using the public data in hong kong from january 16 to february 15, 2020. Front. Phys. 8, 347 (2020).

67. Clifford, S. et al. Strategies to reduce the risk of sars-cov-2 re-introduction from international travellers. medRxiv (2020). https://www.medrxiv.org/content/early/2020/07/25/2020.07.24.20161281.full.pdf.

68. Wölfel, R. et al. Virological assessment of hospitalized patients with COVID-2019. Nature 581, 465–469 (2020).

69. Bullard, J. et al. Predicting infectious severe acute respiratory syndrome coronavirus 2 from diagnostic samples. Clinical Infections Diseases 71, 2663–2666 (2020).

70. ISARIC: International severe acute respiratory and emerging infections consortium COVID-19 report: 13 july 2020. https://media.tdnion.gov.medialibrary/2020/07/ISARIC_Data_Version_COVID-19_Report_13JUL20.pdf.

71. Dreher, M. et al. The characteristics of 50 hospitalized covid-19 patients with and without ards. Deutsches Arzteblatt Int. 117, 271 (2020).

72. Rees, E. M. et al. Covid-19 length of hospital stay: a systematic review and data synthesis. BMC Med. 18, 1–22 (2020).

73. Khalili, M. et al. Epidemiological characteristics of covid-19: a systematic review and meta-analysis. Epidemiol. Infect. 148 (2020).

74. Grasselli, G. et al. Baseline characteristics and outcomes of 1591 patients infected with sars-cov-2 admitted to icus of the lombardy region, italy. JAMA 323, 1574–1581 (2020).

75. Wu, C. et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in wuhan, china. JAMA Intern. Med. 180, 934–943 (2020).

76. Armstrong, R., Kane, A. & Cook, T. Outcomes from intensive care in patients with covid-19: a systematic review and meta-analysis of observational studies. Anaesthesia 75, 1340–1349 (2020).

77. Ministero della Salute. https://www.dati.salute.gov.it/dati/homeDataset.jsp (2019).

Acknowledgements

Our sincere thanks to Dr. Luciano D’Alfonso, Senatore della Repubblica Italiana, Dr. Angelo Borrelli, capo del Dipartimento della Protezione Civile Italiana e commissario per l’emergenza Covid-19, Dr.ssa Eliana Mazzaro, Dott. della Protezione Civile Italiana, Dott. Stefano Marro, Prevenzione Sanitaria presso Ministero della Salute, for responding to our requests in time and for sharing data of hospital and ICU capacities. We thank Rebecca Ludwig for thoughtful discussion and comments on the manuscript. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003480 and the Initiative and Networking Fund of the Helmholtz Association. This project was supported by German Federal Ministry of Education and Research for the project CoViDDec (FKZ: 01KI20102). M.S. and A.B. were supported by the COSMIC Marie Skłodowska-Curie grant (765158). TM and SK were supported by the Helmholtz-Gemeinschaft, Zukunftsthema “Immunology and Inflammation” (ZT-0027), from 2021: (ZT-1-0010). The funding bodies had no role in the design of the study, collection, analysis, and interpretation of the results, or writing the manuscript.

Author contributions

A.B. and M.S. acquired and pre-processed the data, coded the simulation, automated the analysis of the data, analyzed simulation results and led the work. T.M. and M.M.H. developed the reference SECIRD model. A.B., M.S., and T.M. designed the capacity model and protocols for parameter estimation of SECIRD models. A.B. and M.S. developed other versions of the SECIRD model. T.M. derived the literature-based parameter sets and analytical expressions of the reproduction numbers of the models. T.M. and S.K. provided insights through discussion. A.B. wrote the paper. A.B., M.M.H., M.S., and S.B. revised the paper. M.M.H. supervised the study.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43856-022-00139-y.

Correspondence and requests for materials should be addressed to Arnab Bandypadhyay, Marta Schips or Michael Meyer-Hermann.

Peer review information Communications Medicine thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/. © The Author(s) 2022