COVID-19 in Italy: targeted testing as a proxy of limited health care facilities and a key to reducing hospitalization rate and the death toll

Arnab Bandyopadhyay[1,◕,*], Marta Schips[1,◕,*], Tanmay Mitra[1], Sahamoddin Khailaie[1], Sebastian Binder[1], Michael Meyer-Hermann[1,2,3*]

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

◕These authors contributed equally to this work.

* **Correspondence:**
mmh@theoretical-biology.de
arnab.bandyopadhyay@theoretical-biology.de
marta.schips@theoretical-biology.de

Abstract

The novel Coronavirus SARS-CoV-2 (CoV) has induced a worldwide pandemic, notably in Italy, one of the worst-hit countries in Europe, which witnessed a death toll unseen in the recent past. There are potentially many factors, such as infections from undetected index cases, early vs late testing strategies, limited health care facilities etc., that might have aggravated the COVID-19 situation in Italy. We developed a COVID-19 specific infection epidemic model composed of susceptible (S), exposed (E), carrier (C), infected (I), recovery (R) and dead (D) (SECIRD), specifically parameterized for Italy to disentangle the impact of these factors and their implications on infection dynamics to help planning an effective control strategy for a possible second wave. Our model discriminates between detected infected and undetected individuals who played a crucial role in the disease spreading and is not well addressed by classical SEIR-like transmission models. We first estimated the number of undetected infections through a Bayesian Markov Chain Monte Carlo (MCMC) framework, which ranges from ~ 7 to ~ 22 fold higher than reported infections, depending upon regions. We exploited this information to evaluate the impact of the undetected component on the evolution of the pandemic and the benefits of an enhanced testing strategy. In high testing regions like Veneto, 18% of all infections resulted in hospitalization, while for Lombardia and Piemonte, it is 25% and 27%, respectively. We investigated the impact of an overwhelmed health care system upon death toll by applying hospital and intensive care unit (ICU) capacities in the SECIRD model, and we estimated a 10% reduction in death in Lombardia, the worst hit region, if a higher number of hospital facilities had been available since the beginning. Adopting a combined strategy of rapid early and targeted testing (~ 10 fold) with increased hospital capacity would help in avoiding bottlenecks affecting the health care system. Our results demonstrate that the early testing would have a strong impact on the overall hospital accessibility and, hence, upon death toll ($\sim 20\%$ to 50% reduction) and could have mitigated the lack of facilities at the crucial middle stage of the epidemic.

Introduction

The COVID-19 outbreak created a worldwide pandemic causing more than 750,000 deaths and over 21 million total cases worldwide as of August 15th [1]. Until now, COVID-19 has killed more people than SARS and MERS. Developing and deploying a safe and effective vaccine is a lengthy process, and thus controlling the severity of coronavirus disease through a vaccine or targeted therapy—though a major medical concern—is not envisioned in the short term. Alternatively, non-pharmaceutical interventions (NPIs) are recommended in order to control the spread of the disease and have been implemented by several governments, despite the huge economic loss predicted by experts which might lead to a global recession [2]. From the experience of Wuhan, strict measures can be effective in reducing the spread of the virus. How much these NPIs help in the containment of the virus spreading is of high political and academic interest and entirely depends upon the objectives of mitigation [3, 4]. Though different models of NPIs and their implementation methods have been proposed, their impact and effectiveness on disease dynamics are under scrutiny and remain a matter of global discussion [3, 5–8]. On the contrary, Singapore and Hong Kong were able to contain the virus by aggressive testing [9], while South Korea adopted a trace, test, treatment strategy [10]. Both strategies are effective and are equally essential but not necessarily economically optimal. Differently, Japan avoided both lockdown and extensive testing by isolating the whole contact clusters and by heavily relying on self-awareness and discipline of the population to avert high-risk situations [11].

The COVID-19 outbreak originated in Wuhan, People's Republic of China, in early December 2019. Within two months, it had erupted and unfolded with tremendous speed in Italy, which became the European epicenter of the disease spreading, forcing the government to impose a lockdown on March 9th. On March 19th, 3405 people had already died in Italy, thereby surpassing China, while 35731 people were diagnosed as COVID-19 positive. This induced Italy to shut down all non-essential businesses on March 21st. Despite the strict measures applied, a total of 32178 symptomatic people were hospitalized by April 8th in Lombardia, accounting for 12976 hospital admissions, followed by Emilia Romagna (4130), Piemonte (3196), and Veneto (1839) [12]. This led to the complete collapse of the health care 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 reasonable time [13]. Even though the state expanded the hospital and intensive care unit (ICU) capacity, it could not prevent a bottleneck of the health care system which may have caused a high number of deaths for a prolonged time period.

There are potentially many reasons that might have aggravated the COVID-19 situation in Italy, in particularly the distinct demographic structure of Italy when compared to other European nations. In 2019, nearly 23% of the Italian population were aged 65 years or older [14]. Moreover, the average household size and the prevalence of three-generation households are larger in Italy compared to Germany [15]. Limited hospital and ICU capacity might also have contributed to the high number of deaths and, hence, the high case fatality rate (CFR) when compared to other countries. One of the major contributions to the infection spreading came from the asymptomatic and mildly symptomatic infected groups, which mostly remained undetected. At the beginning of the pandemic, Italy focused on testing symptomatic patients only, which resulted in a high proportion of positive tests and high CFR compared to other countries [16]. This left in the society a high proportion of undetected cases, which became a major driver of new positive infections because of their unawareness, while the detected symptomatic cases were, in general, systematically isolated. A different study estimates that in Italy the true number of total infections is around 30 fold higher than that reported, while for Germany it is in the range of six fold [17] (data up to March 17th).

The current study focuses on analyzing the impact of the targeted testing and limited hospital and ICU capacity on the evolution of the pandemic in different Italian regions. Estimates of the surge capacity of Italian ICUs are available [13, 18, 19]. Here, we developed a COVID-19 specific

infection epidemic model composed of susceptible (S), exposed (E), carrier (C), infected (I), recovery (R), and dead (D) (SECIRD). In order to evaluate the COVID-19 situation in Italy in a realistic framework, we first estimated the undetected fraction across different regions of Italy. We used this undetected fraction to parameterize the SECIRD model for describing observed infections, hospitalized, ICU and death dynamics. We studied the influence of undetected and NPIs installed on the time-dependent reproduction number, R_{t}. We analyzed the impact of the overwhelmed health care system upon death toll by limiting hospital and ICU bed availability in the SECIRD model and estimated that an extra 10% people died in Lombardia due to such difficult circumstances. We estimated that the impact of a strategy combining targeted testing with home quarantine (~ 10 fold more isolation of infected) and a hospital capacity at its current elevated state would have significantly reduced the death toll by circa 20% to 50%.

Methods

Data preprocessing

Demographic and mortality data are available from the Italian Institute of Statistics' (ISTAT) website [20, 21]. ISTAT collects mortality data from the Italian *National register office for the resident population* (ANPR). Daily deaths from 2015 to 2019 are available stratified by gender and age. For 2020, data are available up to April 15th, listing 6866 municipalities and covering 86% of the Italian population. For each region, the municipalities listed include those for which the death data set is complete and we, accordingly, extracted demographic data for these municipalities. Here, we considered seven age groups: 0-20, 21-40, 41-50, 51-60, 61-70, 71-80 and 81+. For each region and for each age range, we counted the daily deaths for the period January to April 15th 2015-2020. We then summed up daily deaths of municipalities by age groups. Similarly, from the demographic data we calculated the population size of each age group. Due to the unavailability of demographic data for 2020, we used 2019 data as a proxy of 2020. To calculate the Infection Fatality Rate (IFR) for Italy, we considered mortality and demographic data of municipalities across all regions for which data are complete and followed the same procedure.

Bayesian estimation of COVID-19 IFR

To estimate IFR, we implemented maximum likelihood estimation for the binomial distribution of the mortality rate under a Bayesian framework. Our model is adopted from [22] and modified accordingly. The model assumes that the observed number of deaths in the COVID-19 period (February 20th to April 15th) of each year is binomially distributed according to:

$$
\begin{aligned}
D_{\mathrm{a,y}} &\sim \mathrm{Binomial}(\delta_{\mathrm{a}}, N_{\mathrm{a,y}}) && \text{for y } \in [2015, 2019] \\
D_{\mathrm{a,2020}} &\sim \mathrm{Binomial}(\delta_{\mathrm{a}} + \delta_{\mathrm{a}}^{\mathrm{Covid}} \cdot \theta, N_{\mathrm{a,2019}}) && \text{for y } = 2020
\end{aligned}
$$

where a denotes the seven age groups, mentioned in the previous section and y denotes the year. $D_{\mathrm{a,y}}$ and $N_{\mathrm{a,y}}$ denote total death and population of age a in year y, respectively. δ_{a} is the baseline death rate of age a and is heterogeneous across age groups. To model deaths in 2020, in addition to δ_{a}, we considered a COVID-19 death rate, $\delta_{\mathrm{a}}^{\mathrm{Covid}}$ multiplied with the exposed fraction, θ. $\delta_{\mathrm{a}}^{\mathrm{Covid}}$ is the IFR of age a and it is assumed to be absent in the previous years. $\delta_{\mathrm{a}}^{\mathrm{Covid}}$ is sampled from a uniform distribution with range between 0 and 0.2. The upper bound has been chosen considering that the IFR, by definition, is equal or less than the CFR. For COVID-19, the reported CFR is within the range of 5% to 15%. θ is the infection rate (IR) or attack rate and denotes the fraction of the population that is exposed. As supported by a recent seroprevalence

study [23] it is assumed to be constant across all age groups. Moreover, seroprevalence studies indicate that population-wide immunity is, in general, less than 50% [23–25], thus we sampled θ from a beta distribution for which the density peaks between 20% and 40%. In a nutshell, we are estimating fifteen parameters (considering seven age groups), δ_{a}, $\delta_{\mathrm{a}}^{\mathrm{Covid}}$, and θ from the observation of death data of previous years (2015-2020) classified by age groups (42 data points) given the age distribution of population (2015-2020, 42 values).

We used the following priors to estimate the δ_{a}, $\delta_{\mathrm{a}}^{\mathrm{Covid}}$ and θ:

$$\begin{aligned}\delta_{\mathrm{a}} &\sim \mathrm{Uniform}(0, 0.1)\\ \delta_{\mathrm{a}}^{\mathrm{Covid}} &\sim \mathrm{Uniform}(0, 0.2)\\ \theta &\sim \mathrm{Beta}(3, 5)\end{aligned}$$

For each region, the model was evaluated using the Markov Chain Monte Carlo (MCMC) sampling method. We used 30 independent chains; each drew 50,000 samples from the joint posterior distribution. We discarded the first 5000 as burn-in periods from each chain. We merged the results of all 30 independent chains and calculated the 95% credible interval by using the 95th quantile of the posterior distribution. Trace plots of MCMC and posterior distribution are available in the Supplementary Material (Fig. S3). The total number of infections for each region was calculated from the definition of IFR, which is death over total infection. Here we considered the number of deaths as the excess deaths of February 20th to April 15th with respect to the average of the previous years (2015-2019). Simulation results are reproducible and were carried out using R version 3.6.2 [26–29].

SECIRD-model and basic reproduction number

To understand the impact of potential aggravating factors, namely infections from undetected index cases, early vs late testing strategy, and limited heath care facilities on disease progression, we adapted a COVID-19 specific SECIRD-model from [30, 31] and specifically parameterized it for Italy. The SECIRD-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 asymptomatic (C_{R}) and pre-symptomatic infected (C_{I}), determined as α and $(1-\alpha)$ portion of the exposed, respectively. The pre-symptomatic infected are categorized into detected symptomatic (I_{H} and I_{R}) and undetected mild-symptomatic (I_{X}), determined as μ and $(1-\mu)$ portion of the carrier (C_{I}). Out of the C_{I}, ρ fraction required hospitalization (I_{H}) and $(1-\rho)$ fraction are symptomatic but recover without hospitalization (I_{R}). Further, compartments for hospitalization (H) and intensive care units (U) were introduced to monitor the load on the healthcare system. Similarly, ϑ and $(1-\vartheta)$ represent the fraction of H that requires ICU (H_{U}) or recovered from hospital (H_{R}), respectively. δ and $(1-\delta)$ represent the fraction of ICU who subsequently die (U_{D}) or recover (U_{R}). The recovered compartment (R) consists of recovered patients from different states of the infection. The model is summarized in *SECIRD Model (Reference)* Fig. 1, with parameters in Table S1. The model equations are outlined in the Supplementary Material.

In order to understand the the number of secondary infections per primary case, we measured time-varying reproduction number (R_{t}) [32] and how it is altered by different measures adopted by society, we opted for fitting the data in one week shifting time windows. This approach has two advantages: first, the reproduction number R_{t} is determined as a time-dependent variable and thus it will reflect the impact of NPIs on the infection dynamics; second, the moving-window helps smoothing the evaluation of sudden jumps in the data, which most likely reflects reporting delays. We further fine-tuned this approach by including the history of the viral spreading in each time window, which is stored in the state space of the model in the previous window. In

each time window, a best fit of the model parameters was chosen based on the cost function value (squared difference between data and simulation). In the next window, the state of the model of the previous time window is used as initial condition for the subsequent fitting inside the time window.

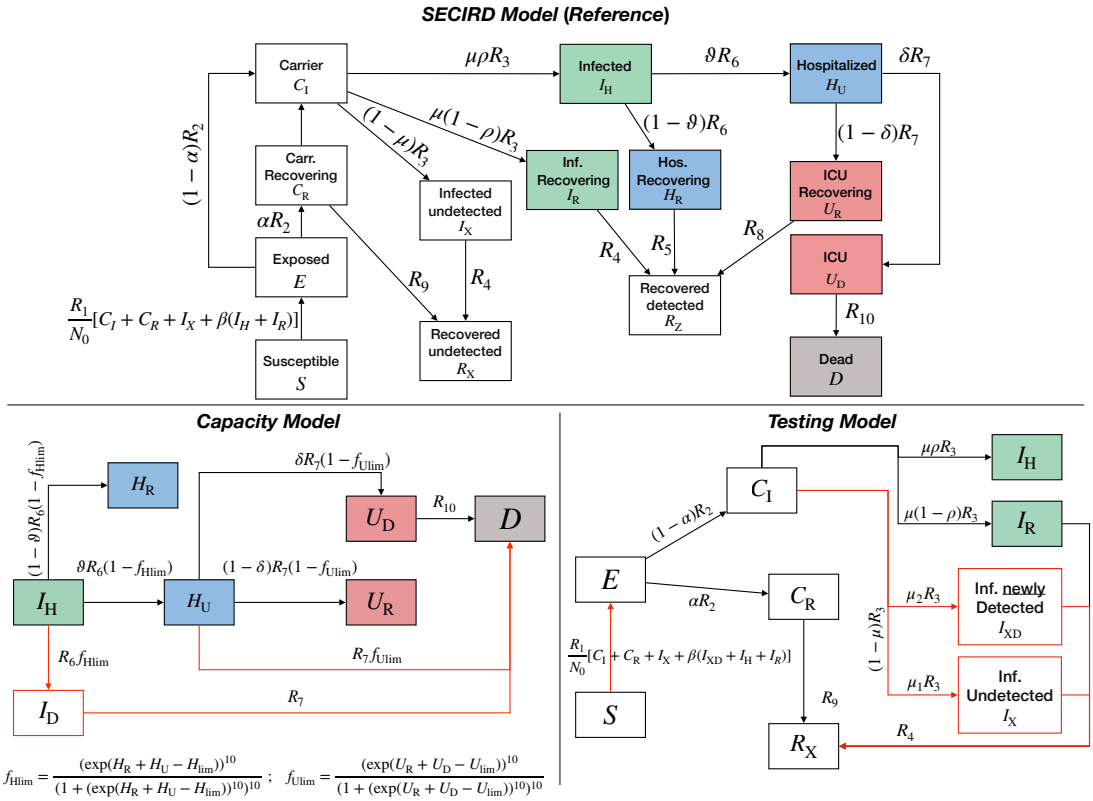

Fig 1: **Schematic of the models.** *SECIRD Model (Reference)*: schematic representation of the SECIRD model used to evaluate the *Reference* and *Asymptomatic* model. The model distinguishes healthy individuals with no immune memory of COVID-19 (susceptible, S), infected individuals without symptoms but not yet infectious (exposed, E) and infected individuals without symptoms who are infectious (carrier, $C_{\mathrm{R,I}}$, asymptomatic and pre-symptomatic respectively), infected ($I_{\mathrm{X,H,R}}$), hospitalized ($H_{\mathrm{U,R}}$) and ICU ($U_{\mathrm{D,R}}$) patients, dead (D) and recovered ($R_{\mathrm{X,Z}}$), who are immune from reinfection. *Testing Model*: modified branch of the SECIRD model, used to evaluate the impact of increasing case detection and isolation on infection dynamics; I_{XD} and I_{X} compartments describe newly detected cases and undetected cases, respectively. *Capacity Model*: modified branch of the SECIRD model to understand the impact of limited hospital and ICU access upon death toll. f_{Hlim} and f_{Ulim} are steep exponential functions diverting the flux from I_{H} and H_{U} to D respectively, when Hospital and ICU reached its current capacity, $H_{\mathrm{lim}}(t)$ and $U_{\mathrm{lim}}(t)$, respectively. Physiological parameters, R_{x} with $x \in [1, \ldots, 10]$ are the transition rates between the different states per day in all panels. Fractional parameters (ρ,ϑ,δ) are subject to the contingent factors, like NPIs installed, self-awareness, availability of hospital beds, etc., and hence are function of time.

The data analysis of the clinical state of all infected cases (up to June 22nd) by ISS showed $\sim 30\%$ asymptomatic cases, with an increasing tendency over time [33–35]. Moreover, in a study performed in Vo' Euganeo, Veneto, the percentage of asymptomatic cases was found to be $\sim 40\%$ [36]. Therefore, we set the asymptomatic fraction $\alpha = 0.4$ in the model. The previously

estimated fraction of undetected cases, $\bar{\mu}$ (see *Estimation of undetected* in Results section, Table 1) is by definition:

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

The R_{t} formula at time window k reads [30]:

$$R_k = R_{1,k}\frac{S_k}{N_k}\left[\frac{1-\alpha}{R_3} + \beta\mu\rho_k\frac{1-\alpha}{R_6} + \frac{\alpha}{R_9} + \beta\mu(1-\rho_k)\frac{1-\alpha}{R_4} + (1-\mu)\frac{1-\alpha}{R_4}\right] ,$$

To generate confidence intervals for R_{t}, we created an ensemble of altered systems by a random modification of the parameters from their reference value such that the total parameter variation, κ, defined as $\log(\kappa) = \sum_{n=1}^{L} |\log\frac{k_n}{k_n^0}|$ [37], is within 20% of its reference value. k^n, k_n^0 and L represent the parameters of the altered system, the reference system and the total number of parameters, respectively. We generated 30 perturbed parameter sets that can quantitatively reproduce the observed data; and all statistical analysis were performed considering these parameter sets. Parameter variations are reported in Fig. S4 and S10.

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_{X}) is detected (I_{XD}) via testing and, hence, contained. The infectivity power of I_{XD} decreases to β ($\beta < 1$) but the infectious period remains unaltered ($1/R_4$) (*Testing Model* in Fig. 1 and Table 2). We introduced a time-dependent undetected fraction μ', which decreases daily by 2%, starting from $\bar{\mu}$ and up to a fixed percentage, 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, $\mu_1(t)$, and the fraction of newly detected cases, $\mu_2(t)$, satisfy the following:

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

with

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

The parameters obtained by fitting the data with the *Reference* model were transferred into the *Testing* model. Transferring parameters maintains the compartmental flow as in the *Reference* model and thus ensures that the result reflects the sole effect of isolating a fraction of undetected infection. Correspondingly, the R_{t} formula is modified as follows:

$$R_k = R_{1,k}\frac{S_k}{N_k}\left[\frac{1-\alpha}{R_3} + \beta\mu\rho_k\frac{1-\alpha}{R_6} + \frac{\alpha}{R_9} + \beta\mu(1-\rho_k)\frac{1-\alpha}{R_4} + \beta\mu_{2,k}\frac{1-\alpha}{R_4} + \mu_{1,k}\frac{1-\alpha}{R_4}\right] ,$$

Capacity model

To estimate the impact of capacity limitations of the health care system, we implemented time-varying capacity constraints on the hospital ($H_{\mathrm{lim}}(t)$) and ICU ($U_{\mathrm{lim}}(t)$) accessibility (*Capacity Model* in Fig. 1 and Table 2) in the SECIRD model, using the available data on the number of hospital and ICU beds in the different regions (Table 3).

Table 3 reports the pre-pandemic capacity and the increased capacity, specifically allocated to COVID-19 patients, with the installation dates. Some regions doubled their capacity and, presumably, this extension of infrastructure has been implemented in a stepwise manner. We assumed a linear increase of the hospital ($H_{\mathrm{lim}}(t)$) and ICU ($U_{\mathrm{lim}}(t)$) capacity from three days

prior to their exhaustion until the installation date, so that it reached the maximum allowed capacity on the installation date. This new limit is available thereafter. The exhaustion date is determined from the data and refers to the day at which the number of hospitalized and ICU patients overcomes the initial capacity. In addition, it was reported that pre-pandemic, overall, 85% hospital beds and 50% ICU beds are occupied [38]. In the *Capacity model*, the remaining percentage of the pre-pandemic capacity (Table 3) has been considered as the baseline capacity of hospital and ICU beds, i.e. the starting values of $H_{\text{lim}}(t)$ and $U_{\text{lim}}(t)$, respectively.

In the model we introduced the I_{D} compartment, where the flux from I_{H} is directed when the hospital capacity is reached. Access to the hospital or ICU of the patients is driven by sharp exponential functions:

$$f_{\text{Hlim}} = \frac{(\exp(H_{\text{R}} + H_{\text{U}} - H_{\text{lim}}))^{10}}{(1 + (\exp(H_{\text{R}} + H_{\text{U}} - H_{\text{lim}}))^{10})^{10}}$$
$$f_{\text{Ulim}} = \frac{(\exp(U_{\text{R}} + U_{\text{D}} - U_{\text{lim}}))^{10}}{(1 + (\exp(U_{\text{R}} + U_{\text{D}} - U_{\text{lim}}))^{10})^{10}}.$$

$f_{\text{Hlim}}(t)$ drives the flow of the infected population into I_{D} when the hospital capacity is reached, while $f_{\text{Ulim}}(t)$ drives a fatal outcome for patients in need of ICU when the ICU reaches its capacity. We assumed that inaccessibility of hospital or ICU leads to faster death. In particular, when the ICU capacity is reached, people in the hospital compartment (H_{U}) die after $1/R_7$ days which is less than hospital-ICU-dead route ($1/R_7 + 1/R_{10}$). Similarly, when the hospital capacity is reached, people in the infected compartment (I_{H}) die after $1/R_6 + 1/R_7$ days, satisfying $1/R_6 < 1/R_6 + 1/R_7 < 1/R_6 + 1/R_7 + 1/R_{10}$ (see *SECIRD models* in the Supplementary).

The *MaxCap* model is defined by the same equations as the *Capacity* model, but the parameters, U_{lim} and H_{lim} are set at the maximum ICU and hospital limits, respectively, from the beginning of the simulations to represent a health care system not overwhelmed.

Parameterization

The parameters in the model are classified in physiological and fractional parameters. Physiological parameters are those whose value mainly depends on the nature of the virus ($R_{\text{x}}, x = 2, \ldots, 10$). We first determined the range of the physiological parameters, except death rate (R_{10}), based on a literature search [39, 40] (Table S1) and then estimated their value by considering the first two weeks' data points, representative of the exponential growth phase. In the course of the pandemic, factors like NPI installation, change in detection and treatment strategies, public awareness and self-isolation,and hospital accessibility substantially alter the disease dynamics by impacting the transmission probability, the amount of undetected infected in the system and the death rate. Fractional parameters (ρ, ϑ, δ), together with R_1 and R_{10}, are the parameters affected by these contingent factors. The parameters α and μ have been previously discussed (see *SECIRD-model and basic reproduction number* section in Methods).

We estimated model parameters by minimizing the sum of squared differences between the observed data (infected, hospitalized, ICU patients and death numbers (Italy Data on Coronavirus 2020 [41])) and model simulation using Matlab's nonlinear least-squares optimizer. The method was repeated separately for each region in Italy. We opted for a moving time window with the size of 7 days to account for temporal changes in disease transmission. The values of R_{x}, ($x = 2, \ldots, 9$) were kept constant throughout the window fitting. Fractional parameters (ρ, ϑ, δ), R_1 and R_{10} were fitted in each window.

An automated method was implemented and parameter estimation carried out in Matlab 2019b [42] with a combination of the Data2Dynamics framework [43].

Results

Estimation of undetected cases

The severity of an epidemic can be characterized by the case fatality rate (CFR), defined as the percentage of deaths among the total number of diagnosed infections, and has been of high interest since the very beginning of the COVID-19 outbreak [44, 45]. On the other hand, infection fatality rate (IFR) is defined as the percentage of deaths among all infections, including the undiagnosed infections. For COVID-19, the true number of cases is unknown as a substantial portion of the infections are either asymptomatic or mildly symptomatic and remained undetected [46, 47]. At the beginning of the pandemic, testing was limited only to the symptomatic patients due to clinical findings suggesting that symptomatic cases are the major source of disease spreading [48–52]. It is, therefore, very difficult to get a reliable estimate of the true number of infections. A few methods have already been proposed to quantify the true number of infections and, hence, a realistic IFR [53–55].

To estimate undetected infections amid the COVID-19 pandemic, we analyzed the mortality rate of previous years and deaths in the year 2020. Demographic and death data of the Italian regions have been collected from the Italian Institute of Statistics (ISTAT). According to the report published by ISTAT, from the first COVID-19 death in Italy (February 20th 2020) to March 31th 2020, excess death in 2020 was 25,354 (total 90,946) compared to the previous five years (65,592 as average of 2015-2019). 54% of the additional deaths were diagnosed as COVID-19 positive [56]. In Fig. S1 we show region-wise, weekly deaths from January 1st to April 15th 2020. For the Italian regions where the pandemic started, like Lombardia, Veneto, Piemonte, the observed mortality of this year was substantially higher than previous years. For these regions, we estimated the total infection and associated IFR by implementing a Bayesian framework by adapting a standard binomial model (see Methods and [22]).

After the first identification of a COVID-19 case on February 20th 2020 in Codogno Hospital, near Lodi, Lombardia [19], the number of reported positive cases increased to 36 in the next 24 hours and, interestingly, the new cases were not linked to the first case, suggesting that the virus was circulating before its first identification. This is reflected in our estimation of undetected cases (Table 1). High increases in death over this year were observed in some cities, like Bergamo (568%), Cremona (391%), Lodi (371%), Brescia (291%), Piacenza (264%), Parma (208%) [56]. In the northern regions of Italy, especially where the initial outbreak occurred, for example in Emilia Romagna, Piemonte and Veneto, undetected infections were nearly 10 fold higher than the reported cases; and in Lombardia more than 21 fold. We observed substantial heterogeneity of the IFR across different age groups. For the age ranges below 60, it was determined as low as 0.05%. IFR was substantially higher in the 81+ age group (9.5% to 20%) and very stringent across the regions (Fig. S2). Despite Italy having the highest COVID-19 deaths in Europe, estimated infection rates (IR) were relatively low (highest in Lombardia $\sim$ 13%) across all regions, and hence the population was far from reaching the herd immunity threshold ($\sim$ 70%, assuming no previous immune memory). Our estimated total infections (detected + undetected) are close to the numbers reported in [55, 57].

Correlation between undetected cases and test frequency

Despite being hit hardest by COVID-19, some of the Italian regions handled this crisis situation better and managed to contain the virus. For instance, in Veneto, CFR was 6.4%, 3 fold lower when compared to Lombardia at 18.3%. This is also reflected in the IR, 2.61% in Veneto while in Lombardia it was 13% despite their geographical proximity. The testing strategy implemented by these two regions was completely different. Most regions, like Lombardia and Piemonte, followed the World Health Organization (WHO) and central health authority indications by mainly testing the symptomatic cases, while Veneto implemented a much more extensive population testing. Toscana followed a testing strategy very similar to Veneto by ramping up its testing capacity

Table 1: Estimation of the total number of infections, the Infection Rate (IR), the Infection fatality rate (IFR) through maximum likelihood in a Bayesian framework based on the data provided by ISTAT up to April 15th [20, 21]. Age specific IFR is reported in Fig. S2.

Areas	IFR in % (95% CI)	Estimated Infections (Undetected %)	IR in % (95% CI)	CFR in %	Reported Infections
Italy	1.58 (1.04 - 1.84)	2627807 (93.73%)	4.37 (3.8 - 6.64)	13.11	165155
Emilia Romagna	1.84 (1.03 - 2.24)	252985 (91.69%)	5.79 (4.84 - 10.22)	13.26	21029
Liguria	2.08 (1.15 - 2.6)	85924 (93.09%)	5.63 (4.57 - 10.01)	13.6	5936
Lombardia	1.66 (1.03 - 1.9)	1390759 (95.53%)	13.83 (12.16 - 22.19)	18.3	62153
Marche	1.88 (0.88 - 2.47)	58555 (90.62%)	3.93 (3.05 - 8.11)	13.56	5503
Piemonte	1.73 (0.78 - 2.12)	258792 (92.94%)	6.1 (5.06 - 13.4)	11.05	18229
Toscana	1.63 (0.69 - 2.36)	62671 (87.77%)	1.43 (0.99 - 3)	7.25	7666
Valle d'Aosta	1.54 (0.73 - 2.34)	9785 (90.19%)	9.74 (6.4 - 17.94)	12.63	958
Veneto	1.3 (0.57 - 1.71)	141466 (89.67%)	2.77 (2.19 - 6.09)	6.43	14624

quickly. To determine whether implementing different testing strategies succeeded in keeping the undetected and the overall infection amount under control, we investigated the association between the rate of tests performed by regions and total infections (Table 1) in the early phase of the pandemic. We measured the Pearson correlation coefficient between the total infections (up to April 15th and including undetected infections) normalized by the population size and the total tests performed (up to April 15th) per reported infection. This yielded a significant correlation with a coefficient of -0.71 (Fig. 2A). Moreover, the number of tests per infection was found to strongly correlate with the CFR/IFR ratio (Fig. 2B), which exceeds 1 when undetected cases are present. Clearly, Emilia Romagna, Piemonte and Lombardia are positioned in the low test high undetected area unlike Veneto and Toscana. Benefits of early testing can also be inferred from the infection dynamics. Regions with a high amount of testing, e.g. Veneto and Toscana, flattened the infection curve by the middle of April, while for Lombardia, Liguria and Piemonte, it was delayed by three weeks (first week of May, Fig. 3A). This shows the benefit of isolating cases by adopting an intensive and early testing strategy.

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 SECIRD *Reference* model to estimate the reproduction number R_{t} [30, 32]. We used this model to explain the dynamics of infected, hospitalized, ICUs and death numbers provided by the *Protezione Civile Italiana* [41] (Fig. 3A). We opted for a moving time window of 7 days to smooth out the reporting delays that might induce artificial fluctuations (see Methods). We then shifted the window by one day and evaluated the current window by using the result of the previous window as the initial state. We followed the same procedure for 30 different parameter sets. Parameter variation for Italy is shown in Fig. S4.

We then used the fitted parameters to obtain the time-varying reproduction number, R_{t} (Fig. 3B). The sudden increase in reported cases resulted in an overshoot in the R_{t} curve at the beginning. As nationwide NPIs were installed, public awareness increased and people started self-isolation and social distancing measures, 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 upon 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}. Sensitivity analysis revealed that R_{t} is highly sensitive towards the changes in R_1 and $\bar{\mu}$ (Fig. S7). The impact of installed NPIs and, hence, R_1 is embedded within the decreasing proclivity of R_{t} across the different regions. The impact of the undetected cases on the R_{t} evolution can as well be inferred through the SECIRD model by comparing the R_{t} of

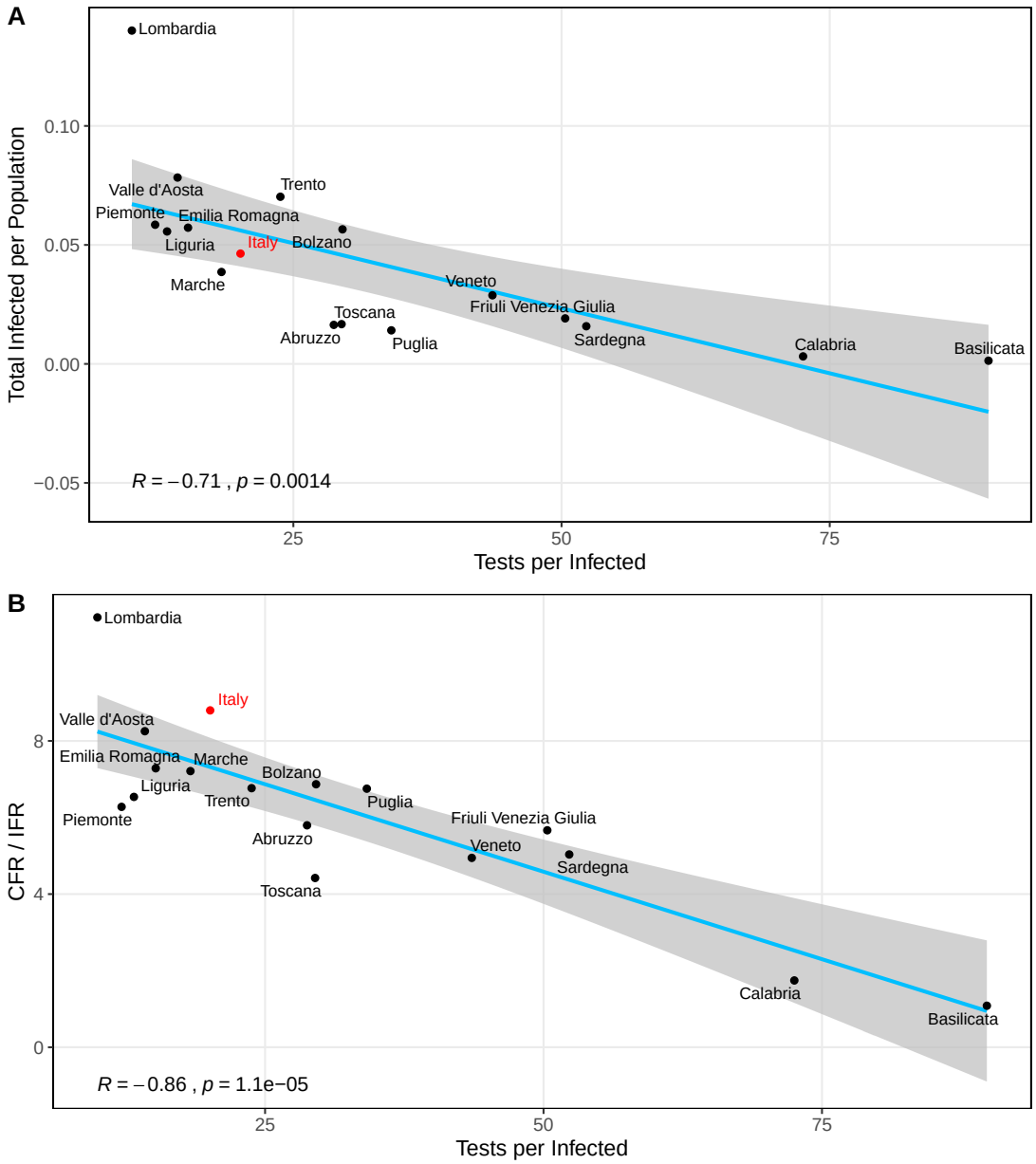

Fig 2: **Impact of the testing frequency on the regional epidemics.** Pearson correlation between: **(A)** number of tests performed per infection and total (detected + undetected) infections normalized by the total population, **(B)** number of tests performed per infection and CFR to IFR ratio. Data are considered up to April 15th. Light blue line: linear regression fit; gray shaded area: standard error; black dots: region-specific values; R: correlation coefficient; p: significance.

the *Reference* model with a model having lower undetected cases. With this aim, we defined the *Asymptomatic* model by imposing $\bar{\mu} = \alpha$, $\mu = 1$.

In the *Asymptomatic* model, all symptomatic infections are detected and, hence, their infectivity is reduced ($\beta < 1$). The lack of the higher infectivity power 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 curves resulting from the *Asymptomatic* are much lower than the R_t curves resulting from the *Reference* model (blue boxplot and red boxplot, respectively, in Fig. 3). On the long term, instead, the R_t curves from the two models converge, despite the fraction of undetected cases being constant throughout the simulations: $\bar{\mu}$ for the *Reference* and α for the *Asymptomatic*. 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 higher in the pre-lockdown period and it vanishes over time. Thus, the influence of undetected infections on R_t wanes (Fig. 3B), most likely as a result of restricted mobility following nationwide lockdown. This illustrates the effectiveness of lockdown in the containment of the disease.

Table 2: Features and purposes of the SECIRD models used in this study. Models' equations are in the Supplementary Material.

SECIRD models	Features	Purpose
Reference	$\bar{\mu}$ inferred from IFR $\alpha = 0.4$ $\mu = (1-\bar{\mu})/(1-\alpha)$	Fit the data Parameter estimation R_t evolution
Asymptomatic	$\bar{\mu} = \alpha$	R_t evolution to be compared with the *Reference* model
Testing	*Reference* model features Sequential decrease of undetected percentage μ $\mu_1 = (\mu' - \alpha)/(1-\alpha)$ $\mu_2 = (\bar{\mu} - \mu')/(1-\alpha)$	Quantify the effect on hospital rates and death toll of reduced amount of undetected fraction
Capacity	*Reference* model features Step-wise increase of H_{lim} and U_{lim}	Parameter estimation Impact of limited hospital and ICU capacities
MaxCap	*Reference* model features H_{lim} and U_{lim} fixed at their maximum Parameters fixed as resulting from the *Capacity* model	Estimate the difference in the dead numbers with respect to the *Capacity* model
TestCap	Combined *Testing* and *Capacity* models	Quantify the effect of testing strategy with a higher hospital and ICU capacities

Track, test, and isolate (TTI) to reduce hospitalizations

As the impact of undetected cases on R_t faded with time, we aimed at identifying and quantifying the benefit of track, test, and isolate (TTI) at early and late phases of the pandemic. We simulated three hypothetical situations where a TTI strategy begins one week before the lockdown (March 2nd), the day after the lockdown (March 10th), and one week post lockdown (March 15th). In each scenario, we assumed that enhanced testing resulted in a 2% daily decrease of the undetected fraction starting from its estimated value (Table 1) to 60%. In the SECIRD model, we considered that a fraction ($\mu_2(t)$) of symptomatic undetected cases, I_X, is detected (I_{XD}) , and hence is less infectious (β, see Methods, Fig. 1, *Testing* model in Table 2). To ensure that the altered infection dynamics is due to the reduction of undetected cases, we maintained the compartmental flow of the SECIRD model by using the parameter set of the *Reference* model complemented with a sequential decrease of the fraction of undetected cases $\mu'(t)$ starting from the estimated fraction $\bar{\mu}$ (Table 1). This setting will help to elucidate the impact of isolating undetected infections that could be achieved 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 infected people that subsequently require hospital access. The early detection scenario showed up to 27% reduction in the hospitalized compartment, which reduced death numbers by up to 41% (Fig. 4A-B) depending upon the region and how early testing starts. The increase in the total infected is due to the extra detection of undetected infections and not new infections. This can also be recognized in R_t (Fig. S8): though the total number of infections is drastically higher than the actual reported infections, R_t falls sharply and for Italy it reaches unity three weeks earlier. This result implies that track, test and isolation of suspected positives can be an effective strategy to substantially decrease disease

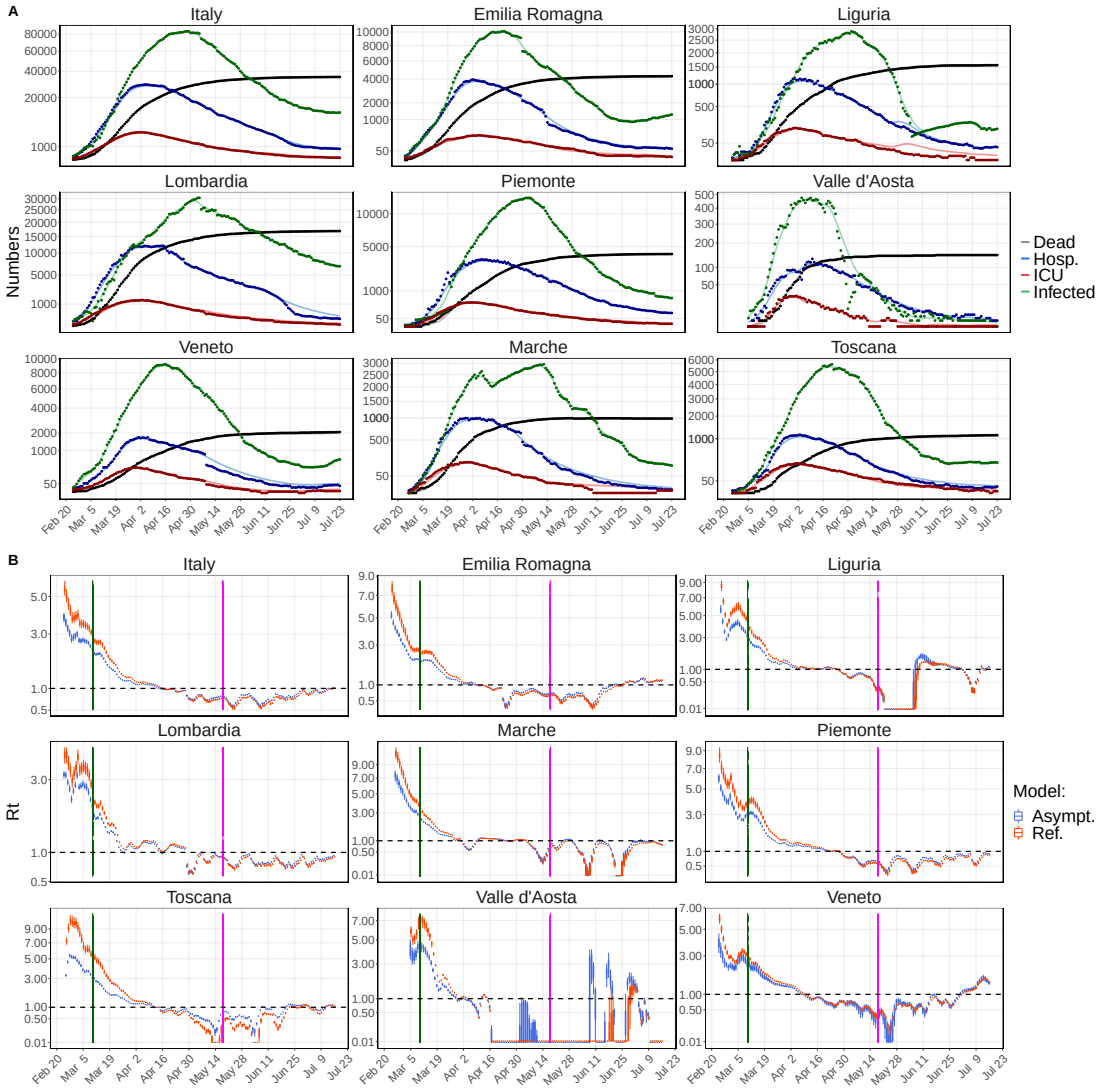

Fig 3: ***Reference* model and time-dependent reproduction number R_t in different regions of Italy. (A)** Infected, hospitalized, ICU and death data (nationwide and region-wise) were fitted in a sliding one week time window. Parameter ranges from Table S1 were used and the transmission rate, R_1, death rate, R_{10} and fractional parameters (ρ, ϑ, δ) were varied in each time window (see Methods); dots: data; continuous lines: sample simulation result out of 30 perturbed parameter sets. **(B)** Boxplot of 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 30 perturbed parameter sets (see Methods) and the box plot shows the median, 25- and 75-percentiles as well as the minimum and the maximum values. Vertical lines correspond to the Lockdown imposition (green) and release (magenta). Black dotted horizontal line represents $R_t = 1$.

propagation. South Korea mounted a targeted testing strategy to contain the disease spreading without imposing strict measures, like lockdown, immigration control etc [10].

Having established from the model that increasing detection and isolation lowers hospitalization rates, we investigated whether the same relation can be inferred from the data. We

calculated the correlation between the median hospitalization rate, defined as the ratio of the sum of the hospital plus ICU patients to the total reported cases, and the total number of tests performed *per capita* up to May 21st, 2020 (Fig. 4C). This data analysis confirms a lower hospitalization rate in high test performing regions. High testing regions, like Veneto or Friuli, exhibited hospitalization rates of 15% and 7%, respectively, in contrast to Lombardia and Piemonte with 25% and 27%. Of note, Veneto performed about twice more tests per head than Lombardia which reduced new cases from undetected index cases and subsequently their hospital access, and was sufficient to induce the significant drop in the hospitalization rate. This qualitative behavior indicates that testing substantially confined the COVID contagion by undetected cases. Our *in silico* experiment shows the cumulative effect of decreasing undetected cases, which is seen in the decreased peak of hospitalization depending upon how early testing begins.

Capacity model and excess dead due to the shortage of hospital beds

Table 3: Hospital and ICU capacity at present and before the pandemic [38, 58]. ICU beds and hosp. beds represent pre-pandemic total beds. In the simulation, we used 50% and 15% of ICU and hosp. beds, respectively. ICU installed and Hosp. installed represent increased allocation specifically for COVID-19 patients. Date ICU and hosp. represent the installed date of newly allocated beds for COVID-19 patients. AP: autonomous province.

Regions	ICU beds	hosp. beds	ICU installed	Date ICU	Hosp. installed	Date hosp.
Abruzzo	109	4410	67	31/03/2020	537	23/04/2020
Basilicata	49	1861	24	17/03/2020	139	17/03/2020
Calabria	153	5739	60	11/04/2020	126	11/04/2020
Campania	506	17977	104	11/04/2020	773	14/04/2020
Emilia Romagna	449	17295	259	24/03/2020	2189	24/03/2020
Friuli Venezia Giulia	127	4333	102	02/04/2020	358	08/05/2020
Lazio	557	20817	323	24/04/2020	1527	21/04/2020
Liguria	186	5690	127	07/04/2020	1241	01/04/2020
Lombardia	859	37767	939	03/04/2020	11673	12/04/2020
Marche	115	5183	132	31/03/2020	638	06/04/2020
Molise	31	1225	12	28/03/2020	31	07/04/2020
Piemonte	317	16313	500	08/03/2020	4451	16/04/2020
Puglia	302	12531	297	11/04/2020	1027	26/04/2020
Sardegna	123	5739	40	14/04/2020	92	07/04/2020
Sicilia	392	15821	312	23/04/2020	1632	04/05/2020
Toscana	377	12021	247	06/04/2020	1350	05/04/2020
Umbria	70	3259	35	25/03/2020	131	11/04/2020
Valle d'Aosta	12	481	25	03/04/2020	262	03/04/2020
Veneto	487	17512	331	17/03/2020	1910	17/03/2020
Bolzano (AP)	40	2047	66	16/04/2020	442	03/04/2020
Trento (AP)	32	2113	70	02/04/2020	382	07/04/2020

Besides installing NPIs and recommending social distancing and self-isolation, strengthening the health care system is also an inevitable part of the government response. According to the data published by the Italian Ministry of Health, before the pandemic started Italy had 3.18 beds per 1000 people with an average occupancy of 75-90% [38], while Germany had 8 beds per 1000 people with $\sim$ 77% occupancy [14]. During this crisis, between March 1st and March 11th, 2020, 9-11% infected people were daily admitted to ICU. Out of $\sim$ 5200 ICU beds (pre-pandemic) in Italy, 2500 were already filled by March th. To cope with this critical situation, each region had planned to boost 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 presumably complicated the crisis management.

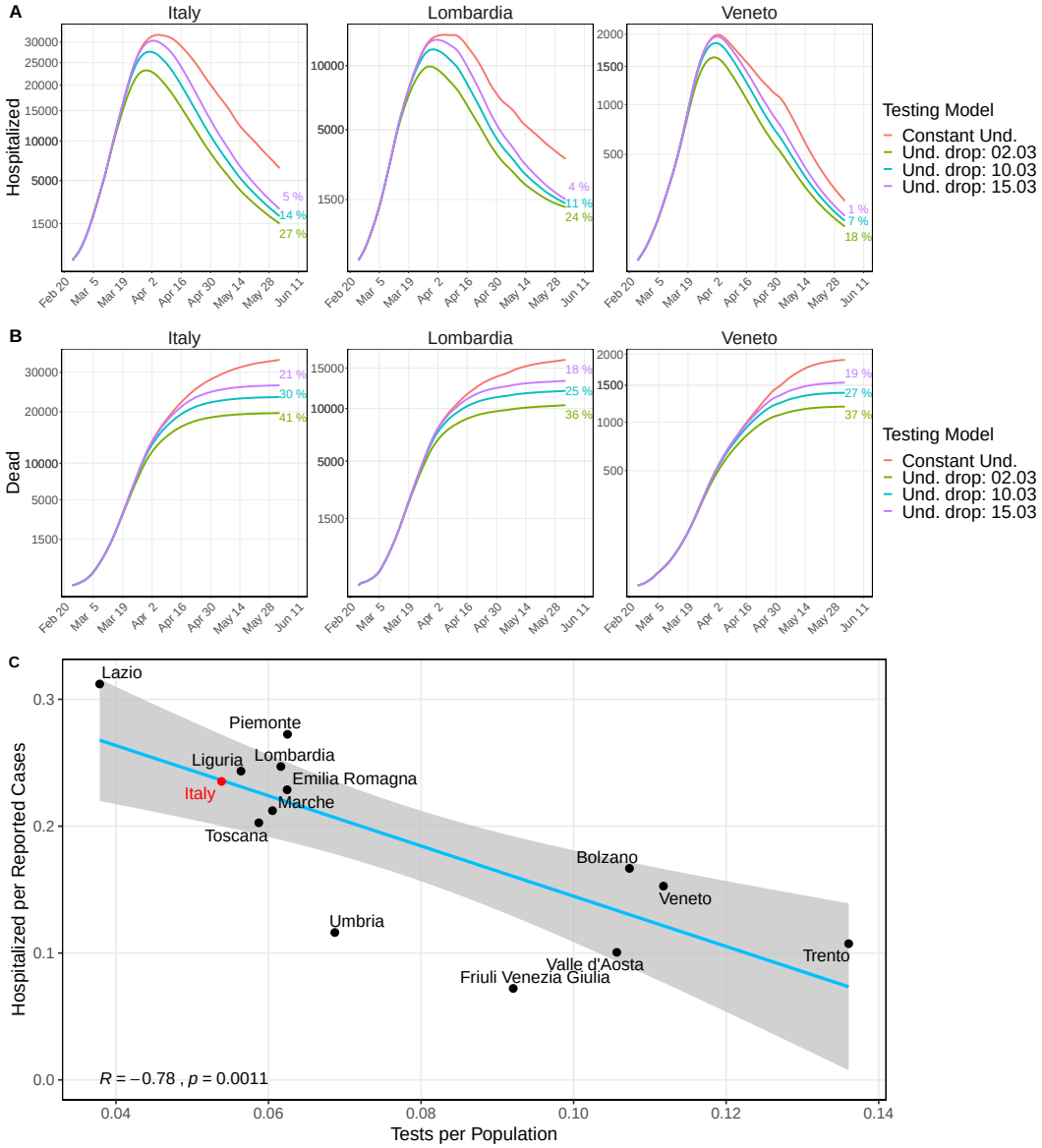

Fig 4: **Impact of testing and isolation on hospitalization and death. (A & B)** Simulation results of the *Testing* model in three evaluated scenarios: undetected cases (Und.) are sequentially decreased from $\bar{\mu}$ to 60% starting one week before the lockdown (green), the day of the lockdown (light blue) and one week post lockdown (lilac). The percentage quantifies the reduction with respect to the fitted *Reference* model (red line). **(A)** Sum of hospitalized and ICU patients; **(B)** number of deaths; **(C)** correlation between number of tests performed per population and the median of hospitalized by reported cases. 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.

Adapting SECIRD, we developed the *Capacity* model by imposing limits on the hospital and ICU accessibility (*Capacity Model* in Fig. 1 and Table 2). Considering the occupancy of hospitals and ICUs (85% and 50%, respectively [38]), the baseline number of hospital and ICU beds was set at 15% and 50%, respectively, of the total number of beds pre-crisis (Initial hosp.

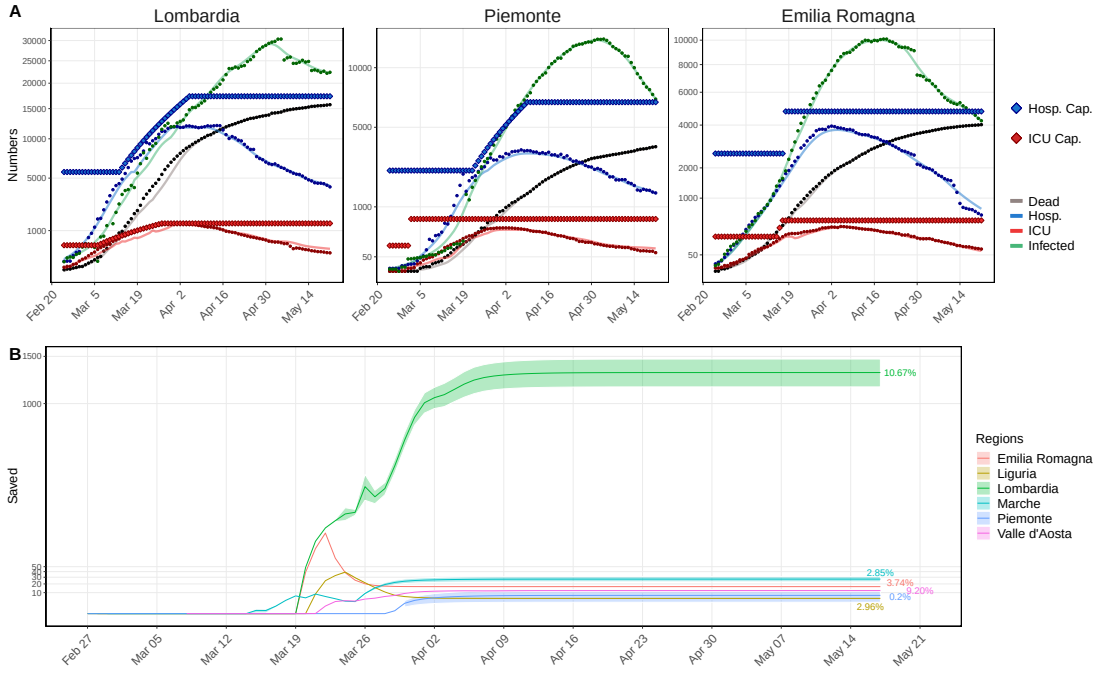

Fig 5: ***Capacity* model and life cost of limited health care system. (A)** Sample *Capacity* model fit of infected (green), hospitalized (light blue), ICU (red), and death (gray) data for the most affected regions. Parameters were fitted as in the *Reference* model. Continuous lines: simulation results; round dots: data; diamonds: linear increase in hospital (blue) and ICU (red) capacity, Table 3. **(B)** Time course of the difference in death numbers between the results of the *Capacity* and the *MaxCap* model. Statistics performed over 30 perturbed parameter sets (see Methods). The reduced death number in the *MaxCap* model with respect to the *Capacity* model is expressed as a percentage.

and Initial ICU in Table 3).

To represent the increase of capacity during the crisis, we used a linear function with a daily increment so that the capacity reaches its target value in the simulation at the same date as in each region (ICU installed and Date and Hosp. installed in Date in Table 3). In addition to imposing the new day-wise capacity onto the SECIRD model, we considered a sharp exponential function (see Methods) to stop the inflow into the hospital and ICU compartments on reaching capacity. We assumed that the unavailability of hospital or ICU beds leads to faster death. We estimated the parameters describing the data with the *Capacity* model using the same protocol applied to estimate parameters for the *Reference* model (Fig. 5). By perturbing the parameters and refitting the data, we generated 30 different parameter sets (Parameter variability is shown in Fig. S10 for Lombardia).

To quantify the impact of the limited hospital capacity on COVID-19-associated deaths, we defined the *MaxCap* model where the hospital and ICU capacities were fixed at 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 of the *MaxCap* model from 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 health care facilities was dramatic in Lombardia with a $\sim 10\%$ difference in the number of deaths, corresponding to ~ 1500 people.

A combined strategy

In a realistic scenario, many regions adopted parallel strategies: they ramped up facilities to test the immediate contacts of an identified infection (tracking and testing, e.g., Veneto [59]); and also planned expansions of health care facilities to accommodate more patients. In light of the reduction in hospitalization and death numbers by extending the track, test, and isolate (TTI) strategy (result of the *Testing* model, Fig. 4A,B) and the reduction in fatalities by increasing the number of hospital facilities (result of the *Capacity* model, Fig. 5B), we aimed at quantifying effects of combining increased hospital capacities with the TTI and considered four scenarios:

1. linear increase of hospital capacity combined with earlier TTI;
2. linear increase of hospital capacity combined with later TTI;
3. maximum hospital capacity from the beginning combined with earlier TTI;
4. maximum hospital capacity from the beginning combined with later TTI.

Thereby, *earlier/later* was assumed one week before/after the lockdown. To simulate scenarios 1 and 2, we transferred the *Capacity* model parameters into the *TestCap* model (Table 2) to keep the compartmental flow intact. To simulate scenarios 3 and 4, we reevaluated the *TestCap* model assuming the hospital and ICU capacity at their maximum levels from the beginning. In all scenarios, improved TTI was simulated as step-wise reduction of the fraction of undetected cases by 2% of the symptomatic undetected cases per day ($\mu'(t)$, *Testing model* in Methods). Given the data, this decrease would be equivalent to $\sim$ 10-fold increased detection.

Fig. 6 reveals several interesting facets. Previously, we had estimated that in Lombardia a 10% reduction of deaths could be achieved by improved hospital infrastructure (Fig. 5B). Boosting hospital capacity with early TTI further reduced death numbers by a significant amount, which ranged from $\sim$ 19% to 51% across different regions (Fig. 6 and S11). Early TTI with $\sim$ 5 fold more testing reduced hospitalization up to $\sim$ 12%, with a subsequent decrease of death toll up to $\sim$ 22% (Fig. S12). This supports the importance of early TTI in controlling the pandemic. Moreover, TTI proved more potent than increasing hospital capacities in limiting the death toll. Even in the late TTI scenario, our simulation shows a decrease in the death number up to 30% (Fig. 6) depending on the region. Of note, adopting an early TTI strategy brings the frequency of hospitalizations down close to the initial hospital capacity before extending the infrastructure (Fig. 6). An exception is Lombardia, where adopting an early TTI strategy led to $\sim$ 18% reduction in the necessity of hospitalizations, while the late TTI strategy did not decrease the simulated peak number of hospitalizations and led to a situation similar to what actually happened. Though the peak size remained unchanged in Lombardia in the late TTI scenario, simulations showed a reduction in the number of deaths of $\sim$ 18%, reaching $\sim$ 26% when combined with an increased hospital capacity.

More testing reduced the frequency of hospitalizations in all regions as is evident from Fig. 4C. These results indicate the importance of early detection and isolation of suspected cases, which substantially reduce the hospitalization rate and could partially compensate for limited health care facilities at the crucial middle stage of the epidemic. On the other hand, despite the strong reduction of hospitalizations and the peak number of deaths, such an early TTI strategy would not suffice to contain the needed number of hospitalizations within the initially available facilities (purple line, Fig. 6). This finding emphasizes the urgent need of extending hospital infrastructure to prevent an overwhelmed health care system.

Herd immunity projection

In order to prepare for a putative second wave of the epidemic, it is helpful to estimate our current standing in terms of immunization of the population. Previously, considering the death data provided by ISTAT up to April 15th 2020, we showed by IFR calculation that the IR was

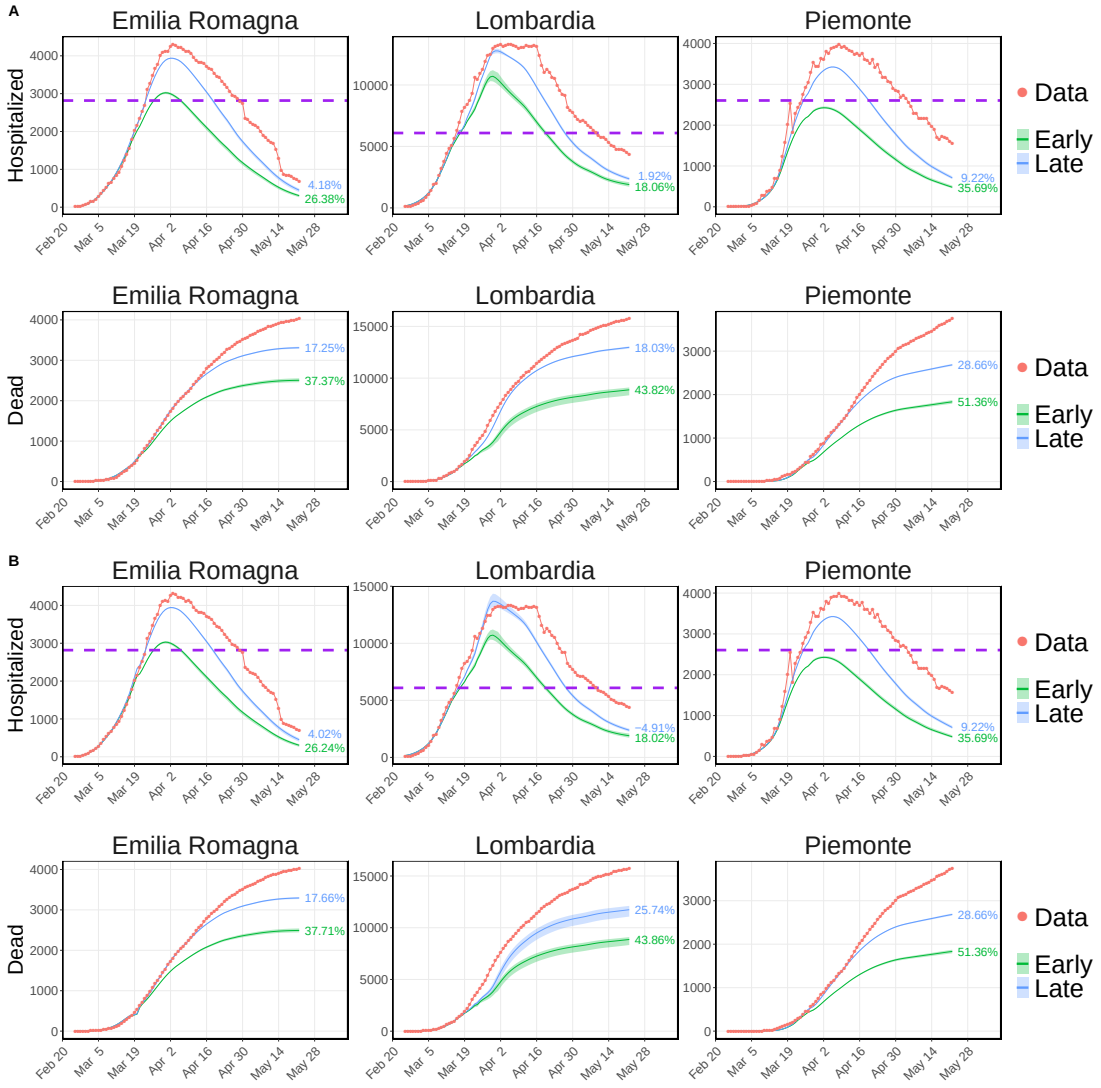

Fig 6: **Impact of combined strategy upon health care.** Simulation results of the *TestCap* model. **(A)** Linear increase of hospital and ICU capacities with either early (green, scenario 1) or late (light blue, scenario 2) testing strategy. **(B)** Maximum capacities from the beginning with early (green, scenario 3) and late (light blue, scenario 4) testing strategy. Statistics were performed by fitting the data with 30 perturbed parameter sets (see Methods) with the *TestCap* model. Line and the shaded region represent the mean and minimum, maximum values, respectively. Hospitalized population is the sum of hospital and ICU patients. The percentage indicates the mean reduction in Hospitalized peak and Death number with respect to the data (red dots). The purple horizontal line is the capacity (hospital + ICU) before the pandemic.

$\sim$ 13% and 10% in Lombardia and Italy, respectively. The fraction of infected people can also be inferred from the SECIRD model. Considering 93% and 95% undetected cases for Italy and Lombardia, respectively, it can be shown from the SECIRD model that the fraction of the people infected up to April 15th, 2020 is close to the range of estimated IR through our Bayesian MCMC framework (vertical black line in Fig. 7). This calibration ensures that a reliable projection of total infected people can be made using the SECIRD model. To this end, we considered two

hypothetical scenarios. First, the undetected fraction is fixed throughout the epidemic. Second, the fraction of undetected cases decreases slowly starting 7 days post lockdown. Considering constant 90% undetected cases, only Lombardia is close to the herd immunity threshold of 70%; in Italy and Veneto only 13% and 10% of the population was infected, respectively. Considering the more realistic scenario where the fraction of undetected cases fades away with time by lockdown measures or public awareness, our model estimated a fraction of 4%, 10%, and 2% of infected people in Italy, Lombardia, and Veneto, respectively. The nationwide predicted total number of infected people is close to the reported number of a recent nationwide seroprevalence study carried out on 64660 people by ISTAT and the Ministry of Health [23].

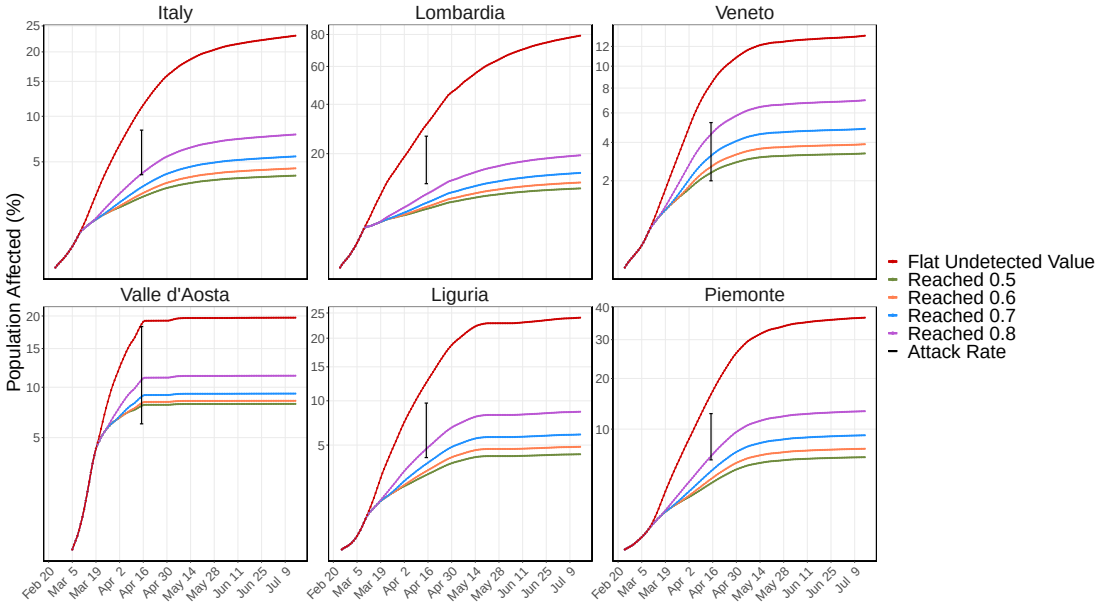

Fig 7: **Herd Immunity projection.** Percentage of the infected population in Italy and in several regions using the fraction of undetected cases constant in time (red) or decreasing over time and reaching 80% (violet), 70% (light blue), 60% (orange), 50% (green). The vertical line represents the estimated infection rate and associated confidence interval up to April 15th, 2020 (Table 1).

Discussion

The COVID-19 pandemic placed a huge social and economic burden on the Italian government and thus confronted policymakers with a dilemma in setting the main objectives of mitigation, e.g., minimizing morbidity and contingent mortality, avoiding an epidemic peak that overwhelms health-care services, minimizing economic loss, and flattening the epidemic curve before a vaccine is made available. Achieving everything by a single intervention is likely to over-promise, especially for the COVID-19 pandemic. One unique feature of COVID-19 is that its serial interval is shorter than its median incubation period and it is shorter than SARS. Thus, substantial secondary infections may occur prior to disease onset [46]. This leads to the rapid spread of the disease by asymptomatic carriers. In order to control the undetected fraction and the spreading through asymptomatic and pre-symptomatic carriers, a rigorous and effective testing strategy is required. The assumption of central health authorities that symptomatic patients are the dominant disease spreaders [60] and the decision of focused testing of symptomatic

patients only has proven to be an insufficient and ineffective strategy as it puts the whole society at risk by reducing the effectiveness of the containment measures imposed at the beginning. Besides implementing NPIs, improving health care facilities is also a part of the government response. Demographically, the Italian population is different than other EU member states, with a substantially older age population, more households containing multiple generations, and a higher occupancy of hospital and ICU beds. COVID-19 is most lethal in older patients and aggravated the health of patients with pre-existing conditions. Moreover, historically, hospital beds per 1000 inhabitants have decreased by 16% since 2010 in Italy, compared to only 3% [14] in Germany. This shortage of hospital beds might also have exacerbated the COVID-19 situation in Italy.

Our retrospective study reveals 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, lockdown alone is not sufficient to control the pandemic and cannot be a long-term plan; neither is it recommended by economists, especially in low and middle-income countries [61]. The power of lockdown lies in restricting the mobility and contacts, reducing the spreading from the mainly asymptomatic undetected infected cases, which results in a general containment of virus spreading. Alternatively, the outcome of the extensive targeted testing strategy adopted by Veneto and Toscana has shown the efficacy in containing the COVID-19 contagion: these two regions contained the undetected fraction and managed to flatten the infection curve one month earlier than Lombardia and Piemonte. Many other countries like Hong Kong, South Korea, Taiwan have successfully flattened the infection curve and controlled regional outbreaks without implementing a strict lockdown. This confirms the effectiveness of a screening strategy and maintaining social distancing. Importantly, we showed that a massive increase in testing significantly reduces the infection transmission and would ultimately reduce the infection to hospitalization rate, which consequently helps easing the pressure on the health care facilities. We showed that ~ 10 fold more testing would release enormous pressure from health care facilities, which in turn significantly reduces the death toll. However, the ~ 10 fold increased testing demands huge testing facilities and is not an economic strategy. Moreover, successful detection depends upon the subject getting tested, and their previous travel and contact history. Alternatively, tracking and targeted testing with home quarantine of possible cases significantly reduce infection transmission and could be adopted instead of rapid mass testing. A more effective strategy would be to put the whole contact cluster into isolation, which Japan adopted [11]. It has to be emphasized that the impact of testing on the spreading of infections is based on the isolation of positively tested individuals. 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.

In the middle of the crisis, many health care 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 testing of health care personnel led to the spread of infection through health care workers. Regular testing of healthcare workers in hospitals and in retirement homes is mandatory to avoid the spread of infection. Improvements in the capacity of hospital and ICU beds is urgent given the aging population [62]. In our model, we have shown that a substantial portion of the death toll could have been prevented by targeted testing and this could have mitigated the effects of limited health care facilities at the crucial middle stage of the pandemic, though it would not have contained the hospitalization within its pre-pandemic limit. Hence, the bottleneck of health care facilities could hardly have been avoided by adopting massive testing alone. We have estimated that we are far from reaching the herd immunity threshold. Thus we believe that the combination of increased detection and isolation of suspected cases, together with improved health care facilities, could reduce mortality in a possible second wave without implementing a strict lockdown.

Acknowledgments

Our sincere thank 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, Dip. 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. The funding bodies had no role in the design of the study, collection, analysis, and interpretation of the results, or writing the manuscript.

Contributions

AB and MS acquired and pre-processed the data, coded the simulation, automated the analysis of the data, analyzed simulation results and led the work. TM and MMH developed the reference SECIRD model. AB, MS and TM designed the capacity model and protocols for parameter estimation of SECIRD models. AB and MS developed other versions of the SECIRD model. TM derived the literature based parameter sets. SB organized the public repository. TM and SK provided insights through discussion. AB wrote the paper. AB, MMH, MS and SB revised the paper. MMH supervised the study.

References

1. worldometers-Coronavirus; 2020. `https://www.worldometers.info/coronavirus/`.

2. COVID-19: Socio-economic impact;. Available from: `https://www.undp.org/content/undp/en/home/coronavirus/socio-economic-impact-of-covid-19.html`.

3. Hollingsworth TD, Klinkenberg D, Heesterbeek H, Anderson RM. Mitigation strategies for pandemic influenza A: balancing conflicting policy objectives. PLoS Comput Biol. 2011;7(2):e1001076. Available from: `https://www.ncbi.nlm.nih.gov/pubmed/21347316`.

4. Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. Imperial College preprint. 2020;.

5. Yang P, Qi J, Zhang S, Bi g, Wang X, Yang Y, et al. The effect of multiple interventions to balance healthcare demand for controlling COVID-19 outbreaks: a modelling study. medRxiv. 2020;p. 2020.05.19.20107326. Available from: `https://www.medrxiv.org/content/medrxiv/early/2020/05/26/2020.05.19.20107326.full.pdf`.

6. Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. Science. 2020;369. Available from: `https://www.ncbi.nlm.nih.gov/pubmed/32414780`.

7. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020;p. 1–5.

8. Dorn F, Khailaie S, Stoeckli M, Binder SC, Lange B, Lautenbacher S, et al. The Common Interests of Health Protection and the Economy: Evidence from Scenario Calculations of COVID-19 Containment Policies. medRxiv. 2020;Available from: https://www.medrxiv.org/content/early/2020/08/15/2020.08.14.20175224.

9. Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? Lancet. 2020;395(10228):931–934. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32164834.

10. Park S, Choi GJ, Ko H. Information Technology–Based Tracing Strategy in Response to COVID-19 in South Korea—Privacy Controversies. JAMA. 2020 06;323(21):2129–2130. Available from: https://doi.org/10.1001/jama.2020.6602.

11. Normile D. Japan ends its COVID-19 state of emergency [Web Page]; 2020. Available from: https://www.sciencemag.org/news/2020/05/japan-ends-its-covid-19-state-emergency.

12. Covid-19 Italy situation report [Web Page]; 2020. Available from: https://covid19.intelworks.io/.

13. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? Lancet. 2020;395(10231):1225–1228. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32178769.

14. Organisation for Ecnomic Co-operation and DEvelopment (OECD); 2020. Available from: https://www.oecd.org/.

15. Iacovou M, Skew A. More than 10% of households in Romania, Latvia and Bulgaria were three-generation in 2008. Statistics in focus. 2011;52:2011. Available from: https://ec.europa.eu/eurostat/documents/3433488/5579620/KS-SF-11-052-EN.PDF/9a190f6f-8949-4d3d-99ad-989479c6b23b.

16. Analysis JHUM [Web Page]; 2020. Available from: https://coronavirus.jhu.edu/data/mortality.

17. Bommer C, Vollmer S. Estimates of the severity of coronavirus disease 2019:a model-based analysis. University of Goettingen [Web Page]; 2020. Available from: https://www.uni-goettingen.de/de/document/download/0af0dcfa623053908de337e1045cf612.pdf/COVID-19%20update.pdf.

18. Charpentier A, Elie R, Laurière M, Tran VC. COVID-19 pandemic control: balancing detection policy and lockdown intervention under ICU sustainability. medRxiv. 2020;p. 2020.05.13.20100842. Available from: https://www.medrxiv.org/content/medrxiv/early/2020/05/22/2020.05.13.20100842.full.pdf.

19. Grasselli G, Pesenti A, Cecconi M. Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast During an Emergency Response. JAMA. 2020;323(16):1545-1546. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32167538.

20. ISTAT; 2020. Available from: https://www.istat.it/it/archivio/240401.

21. ISTAT-popolazione-e-famiglie [Web Page];. Available from: https://www.istat.it/it/popolazione-e-famiglie?dati.

22. Rinaldi G, Paradisi M. An empirical estimate of the infection fatality rate of COVID-19 from the first Italian outbreak. medRxiv. 2020;p. 2020.04.18.20070912. Available from: https://www.medrxiv.org/content/medrxiv/early/2020/05/14/2020.04.18.20070912.full.pdf.

23. Primi risultati dell'indagine di sieroprevalenza sul SARS-CoV-2; 2020. Available from: https://www.istat.it/it/archivio/246156.

24. Knabl L, Mitra T, Kimpel J, Roessler A, Volland A, Walser A, et al. High SARS-CoV-2 Seroprevalence in Children and Adults in the Austrian Ski Resort Ischgl. medRxiv. 2020;Available from: https://www.medrxiv.org/content/early/2020/08/22/2020.08.20.20178533.

25. Arora RK, Joseph A, Van Wyk J, Rocco S, Atmaja A, May E, et al. SeroTracker: a global SARS-CoV-2 seroprevalence dashboard. The Lancet Infectious Diseases. 2020;.

26. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2018. Available from: https://www.R-project.org/.

27. Youngflesh C. MCMCvis: Tools to visualize, manipulate, and summarize MCMC output. Journal of Open Source Software. 2018;3(24):640.

28. Plummer M, Best N, Cowles K, Vines K. CODA: Convergence Diagnosis and Output Analysis for MCMC. R News. 2006;6(1):7–11. Available from: https://journal.r-project.org/archive/.

29. Sturtz S, Ligges U, Gelman A. R2WinBUGS: A Package for Running WinBUGS from R. Journal of Statistical Software. 2005;12(3):1–16. Available from: http://www.jstatsoft.org.

30. Khailaie S, Mitra T, Bandyopadhyay A, Schips M, Mascheroni P, Vanella P, et al. Estimate of the development of the epidemic reproduction number Rt from Coronavirus SARS-CoV-2 case data and implications for political measures based on prognostics. medRxiv. 2020;p. 2020.04.04.20053637. Available from: https://www.medrxiv.org/content/medrxiv/early/2020/04/07/2020.04.04.20053637.full.pdf.

31. SIMM group: Complete and up-to-date analysis of Germany and all federal states.; 2020. Available from: http://secir.theoretical-biology.de/.

32. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am J Epidemiol. 2013;178(9):1505–12. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24043437.

33. Sorveglianza Integrata COVID-19 in Italia. 2020;Available from: https://www.epicentro.iss.it/coronavirus/bollettino/Infografica_22giugno%20ITA.pdf.

34. Sorveglianza integrata COVID-19: archivio. 2020;Available from: https://www.epicentro.iss.it/coronavirus/sars-cov-2-sorveglianza-dati-archivio.

35. Epidemia COVID-19. Aggiornamento nazionale 15 settembre 2020. 2020;Available from: https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_15-settembre-2020.pdf.

36. Lavezzo E, Franchin E, Ciavarella C, Cuomo-Dannenburg G, Barzon L, Del Vecchio C, et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. Nature. 2020;584(7821):425–429.

37. Barkai N, Leibler S. Robustness in simple biochemical networks. Nature. 1997;387(6636):913–7. Available from: https://www.ncbi.nlm.nih.gov/pubmed/9202124.

38. Annuario Statistico del Servizio Sanitario Nazionale; 2017. Available from: http://www.salute.gov.it/imgs/C_17_pubblicazioni_2879_allegato.pdf.

39. RKI Steckbrief; 2020. Available from: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html.

40. L'epidemiologia per la sanità pubblica, I.S.d.S. Characteristics of SARS-CoV-2 patients dying in Italy Report based on available data on June 4th, 2020; 2020. Available from: https://www.epicentro.iss.it/en/coronavirus/bollettino/Report-COVID-2019_4_june_2020.pdf.

41. COVID-19 Italia - Monitoraggio situazione; 2020. Available from: https://github.com/pcm-dpc/COVID-19.

42. MATLAB. version: 9.7.0 (R2019b). Natick, Massachusetts: The MathWorks Inc.; 2019. Available from: https://www.mathworks.com/.

43. Raue A, Steiert B, Schelker M, Kreutz C, Maiwald T, Hass H, et al. Data2Dynamics: a modeling environment tailored to parameter estimation in dynamical systems. Bioinformatics. 2015;31(21):3558–60. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26142188.

44. Baud D, Qi X, Nielsen-Saines K, Musso D, Pomar L, Favre G. Real estimates of mortality following COVID-19 infection. Lancet Infect Dis. 2020;20(7):773. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32171390.

45. Kim DH, Choe YJ, Jeong JY. Understanding and Interpretation of Case Fatality Rate of Coronavirus Disease 2019. J Korean Med Sci. 2020;35(12):e137. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32233163.

46. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. Int J Infect Dis. 2020;93:284–286. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32145466.

47. Zhou X, Li Y, Li T, Zhang W. Follow-up of asymptomatic patients with SARS-CoV-2 infection. Clin Microbiol Infect. 2020;Available from: https://www.ncbi.nlm.nih.gov/pubmed/32234453.

48. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020;395(10223):507–513. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32007143.

49. Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 2020;395(10223):514–523. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31986261.

50. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med. 2020;382(18):1708–1720. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32109013.

51. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497–506. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31986264.

52. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med. 2020;382(13):1199–1207. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31995857.

53. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis. 2020;20(6):669–677. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32240634.

54. Villa M, Myers JF, Turkheimer F. COVID-19: Recovering estimates of the infected fatality rate during an ongoing pandemic through partial data. medRxiv. 2020;Available from: https://www.medrxiv.org/content/early/2020/04/14/2020.04.10.20060764.

55. Modi C, Boehm V, Ferraro S, Stein G, Seljak U. How deadly is COVID-19? A rigorous analysis of excess mortality and age-dependent fatality rates in Italy. medRxiv. 2020;p. 2020.04.15.20067074. Available from: https://www.medrxiv.org/content/medrxiv/early/2020/05/14/2020.04.15.20067074.full.pdf.

56. ISTAT. IMPATTO DELL'EPIDEMIA COVID-19 SULLA MORTALITÀ TOTALE DELLA POPOLAZIONE RESIDENTEPRIMO TRIMESTRE; 2020. Available from: https://www.istat.it/it/files/2020/05/Rapporto_Istat_ISS.pdf.

57. Vollmer M, Mishra S, Juliette H, et al. Using mobility to estimate the transmission intensity of COVID-19 in Italy: a subnational analysis with future scenarios. Imperial College London. 2020;Available from: https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-20-italy/.

58. Ministero della Salute; 2019. Available from: http://www.dati.salute.gov.it/dati/homeDataset.jsp.

59. Bur n. 37 del 20 marzo 2020; Materia: Sanità e igiene pubblica; 2020. Available from: https://bur.regione.veneto.it/BurvServices/pubblica/DettaglioDgr.aspx?id=417291.

60. Coronavirus-Trasmissione, prevenzione e trattamento; 2020. Available from: https://www.epicentro.iss.it/coronavirus/trasmissione-prevenzione-trattamento.

61. Impact of lockdown measures on the informal economy A summary. International Labour Organization; 2020. Available from: https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/briefingnote/wcms_743534.pdf.

62. Phua J, Weng L, Ling L, Egi M, Lim CM, Divatia JV, et al. Intensive care management of coronavirus disease 2019 (COVID-19): challenges and recommendations. Lancet Respir Med. 2020;8(5):506–517. Available from: https://www.ncbi.nlm.nih.gov/pubmed/32272080.

Supporting information

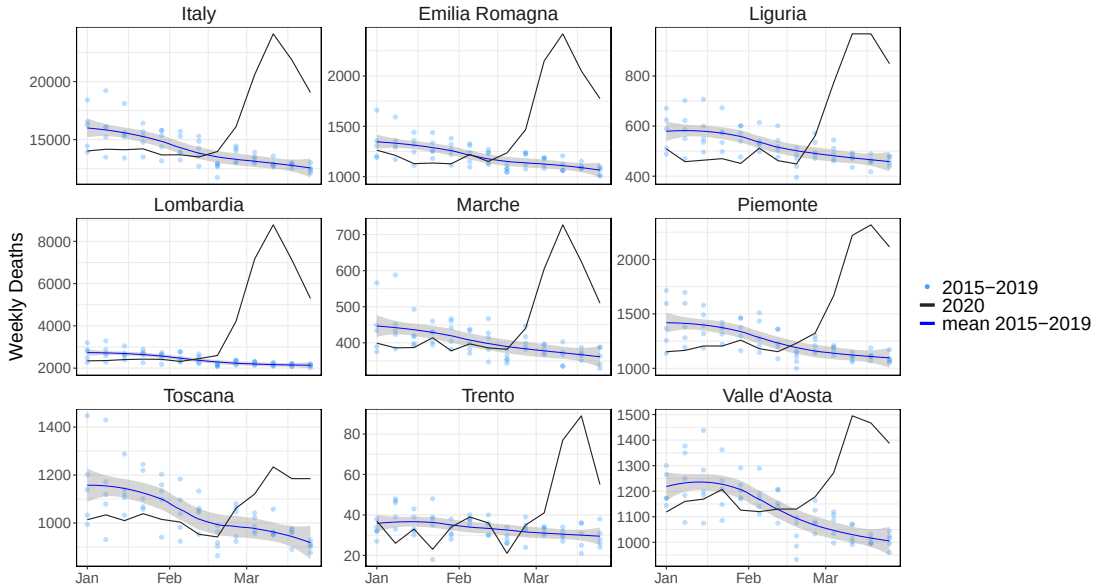

Fig S1: **Excess Deaths.** Weekly deaths for the period of January 1th to April 15th for 2015-2020 of different regions in Italy. Data are acquired from ISTAT and include 6866 municipalities, covering 86% of Italian population. Light blue points: all-cause deaths in 2015-2019; blue line with grey shaded area: spline fit of 2015-2019 deaths; black line: all-cause deaths in 2020.

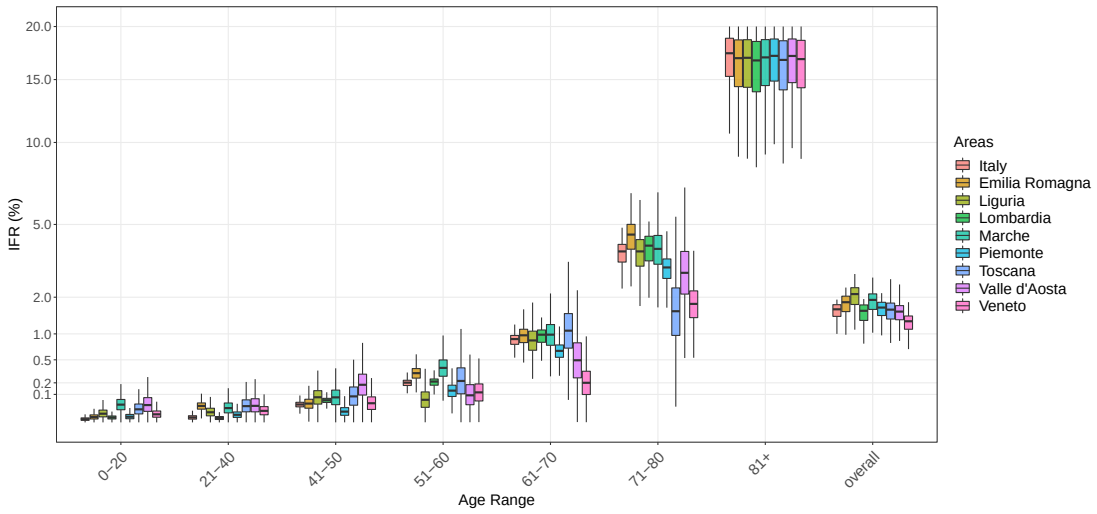

Fig S2: **Age-specific IFR.** Boxplot of estimated IFR across the age ranges and of overall IFR, for Italy and different regions. IFR is estimated by Bayesian MCMC framework (see Methods). IFR is significantly higher in the age group 81+ and is stringent across all regions.

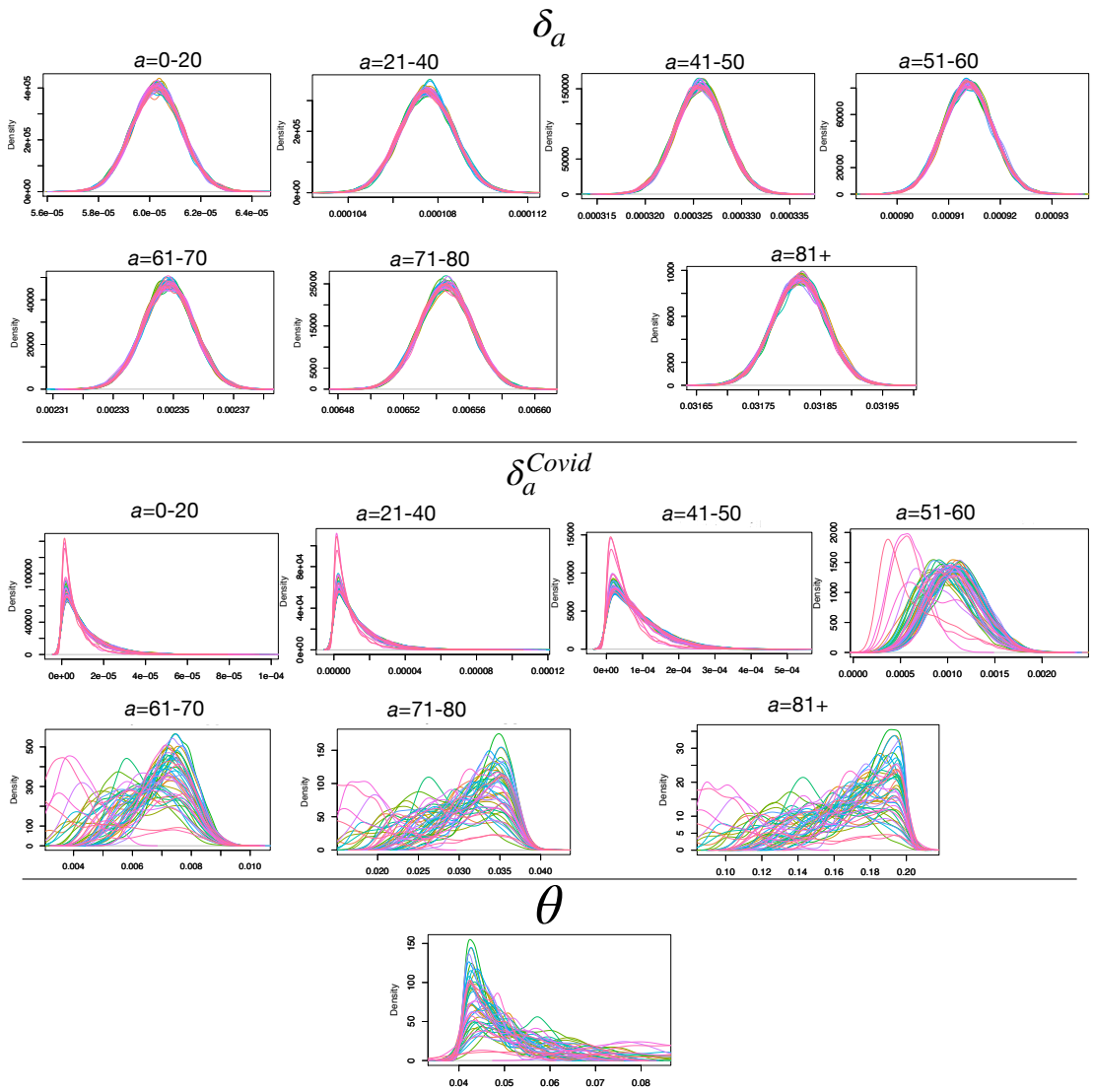

Fig S3: **MCMC results.** Posterior distribution of baseline lethality rates (δ_{a}), IFR ($\delta_{\mathrm{a}}^{\mathrm{Covid}}$) for seven age groups, and Infection Rate (θ) of 30 independent MCMC chains. To estimate δ_{a}, $\delta_{\mathrm{a}}^{\mathrm{Covid}}$, θ, we used following priors: Uniform(0,0.1), Uniform(0,0.2), Beta(3,5). Of note, heterogeneity in $\delta_{\mathrm{a}}^{\mathrm{Covid}}$ distribution is increasing with age.

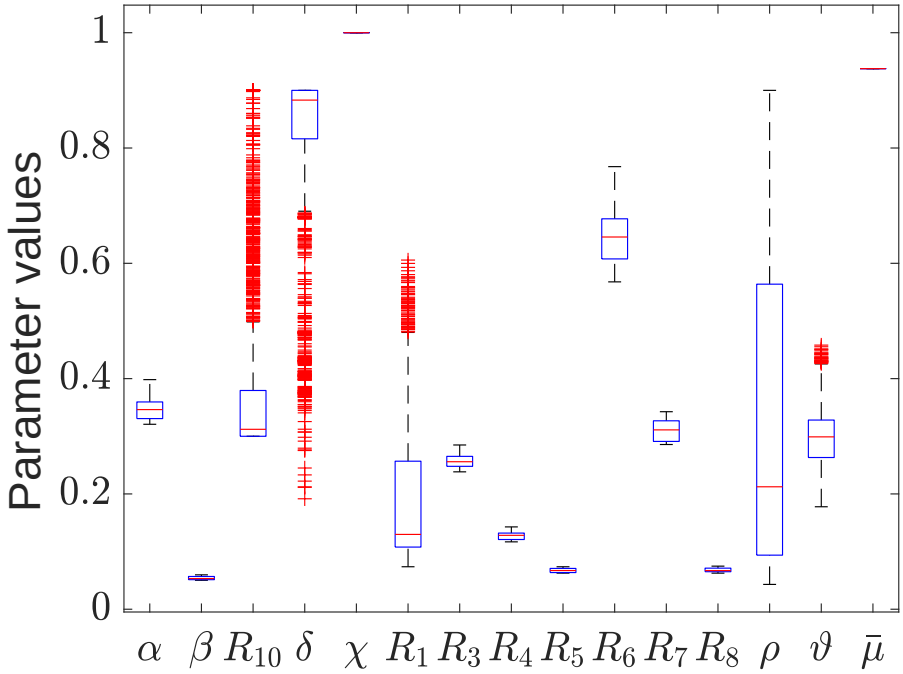

Fig S4: **Parameter variation.** Boxplot of individual parameters as resulting from the fit of Italy data with the *Reference* model by moving window method (see Methods). $R_x, x = 3 - 8$ were estimated by fitting the first 15 data points. 30 parameters sets are generated by perturbing ($\pm 20\%$) this reference set. Remaining parameters were estimated in each window. Statistics are performed over these 30 perturbed parameter sets.

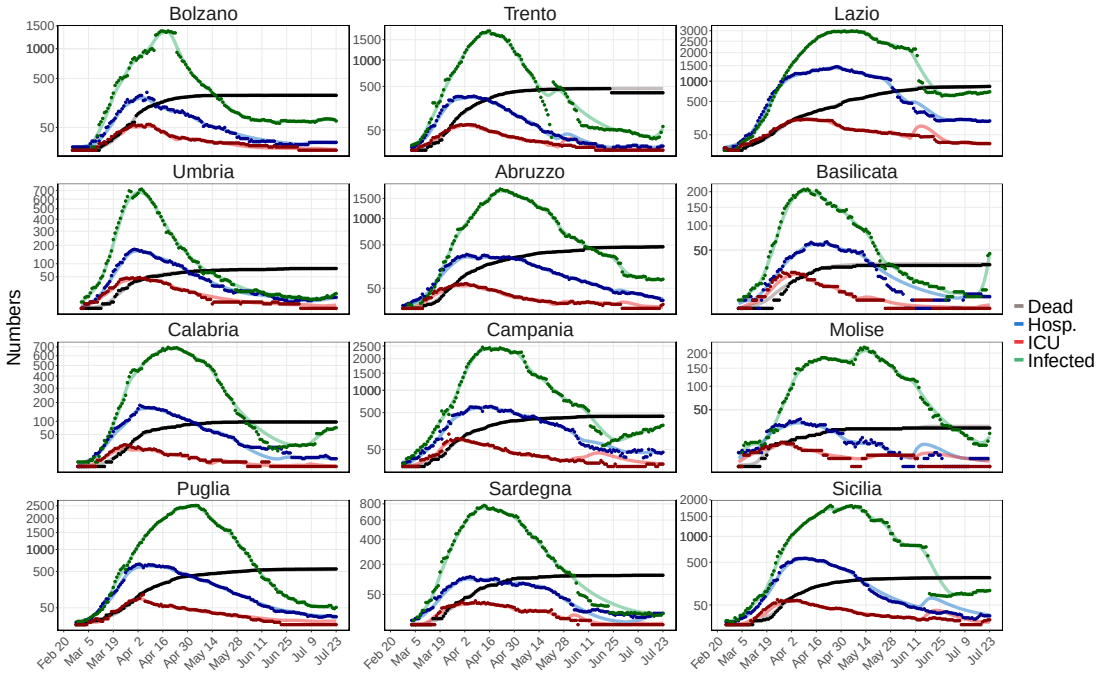

Fig S5: **Fit up to July 23rd** Infected, hospitalized, ICU and death data were fitted in a sliding one week time window. The regions included are complementary to Fig. 3A. Parameter ranges from Table S1 were used and the transmission rate, R_1, death rate, R_{10} and fractional parameters (ρ, ϑ, δ) were varied in each time window (see Methods); dots: data; continuous lines: simulation results.

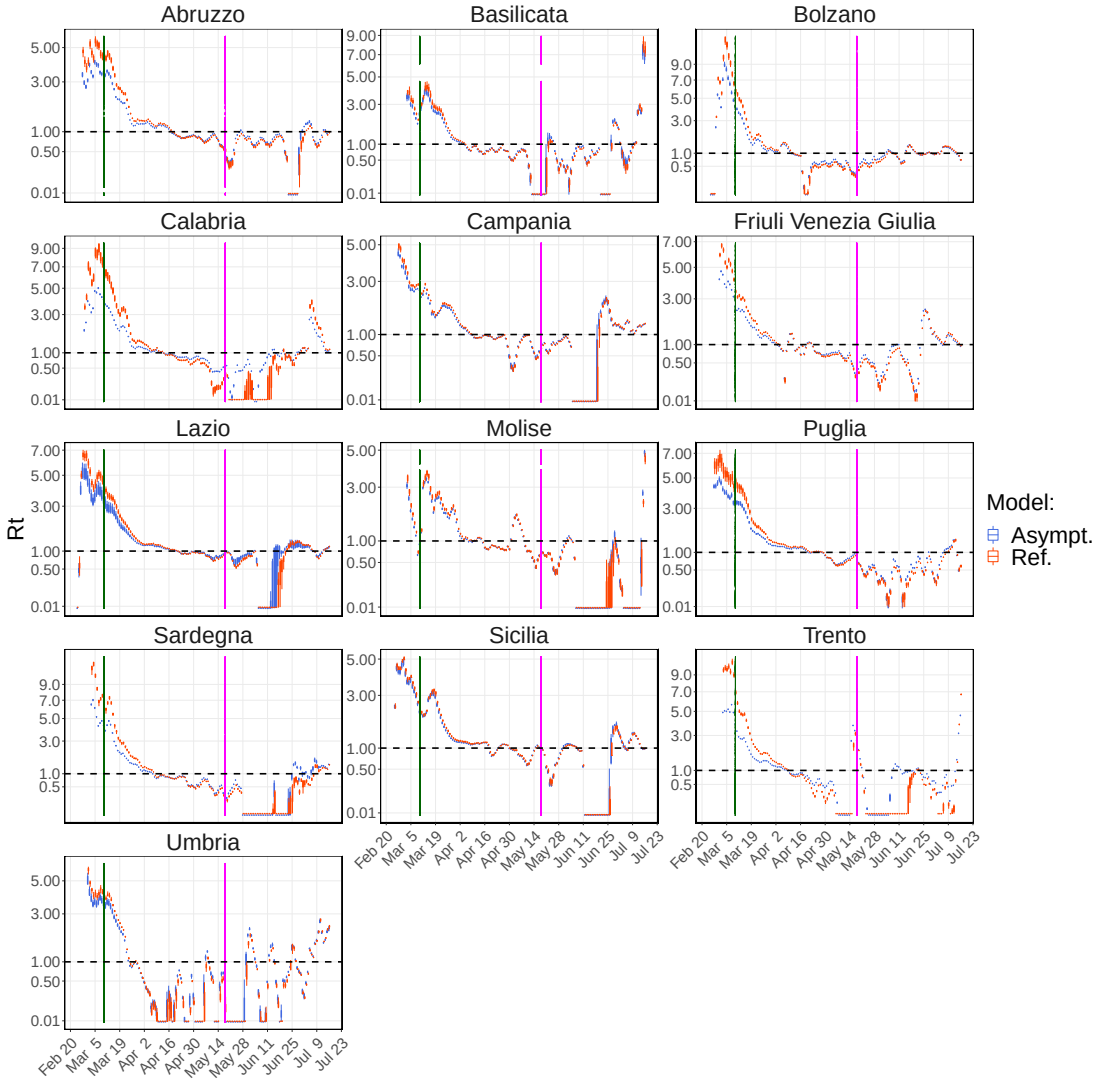

Fig S6: R_t **curves.** Boxplot of R_t results from the fit of *Reference* (red) and *Asymptomatic* (blue) models. The regions included are complementary to Fig. 3B. Statistics performed on the R_t values obtained by fitting the data with 30 perturbed parameter sets (see Methods). Vertical lines correspond to the Lockdown imposition (green) and release (magenta). Black dotted horizontal line represents $R_t = 1$.

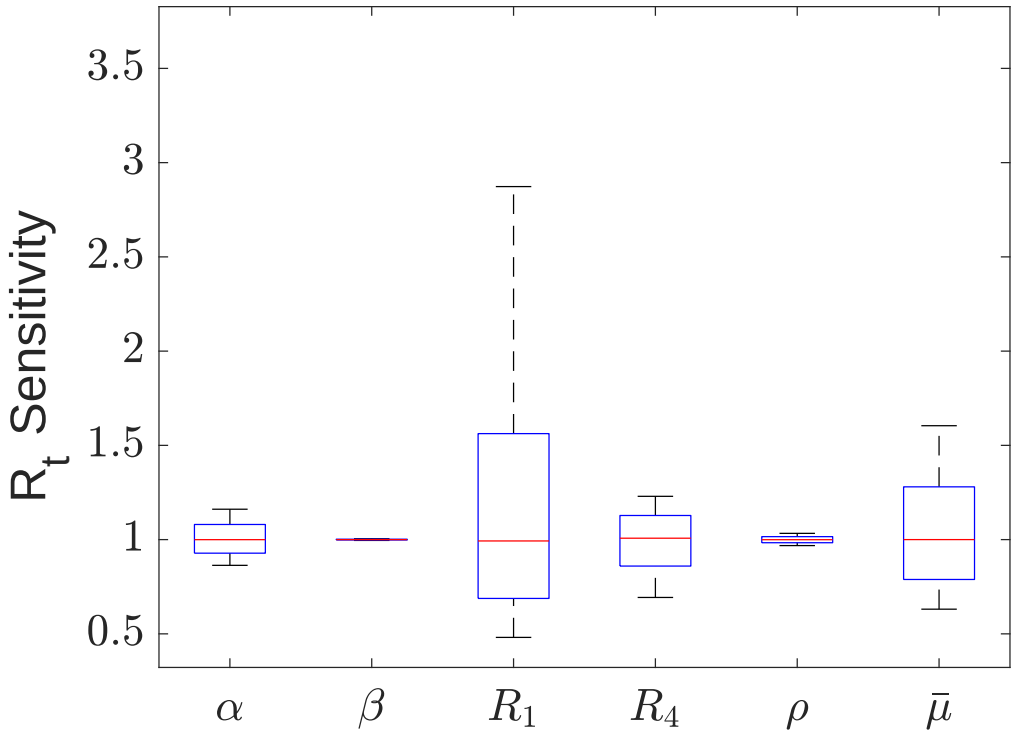

Fig S7: R_{t} **sensitivity.** Sensitivity of R_{t} is performed by perturbing each parameter 30% of its reference value and by simulating the *Reference* model up to three months. Boxplot represents the variation in R_{t} from its reference point. R_{t} is highly sensitive to R_1 and $\bar{\mu}$.

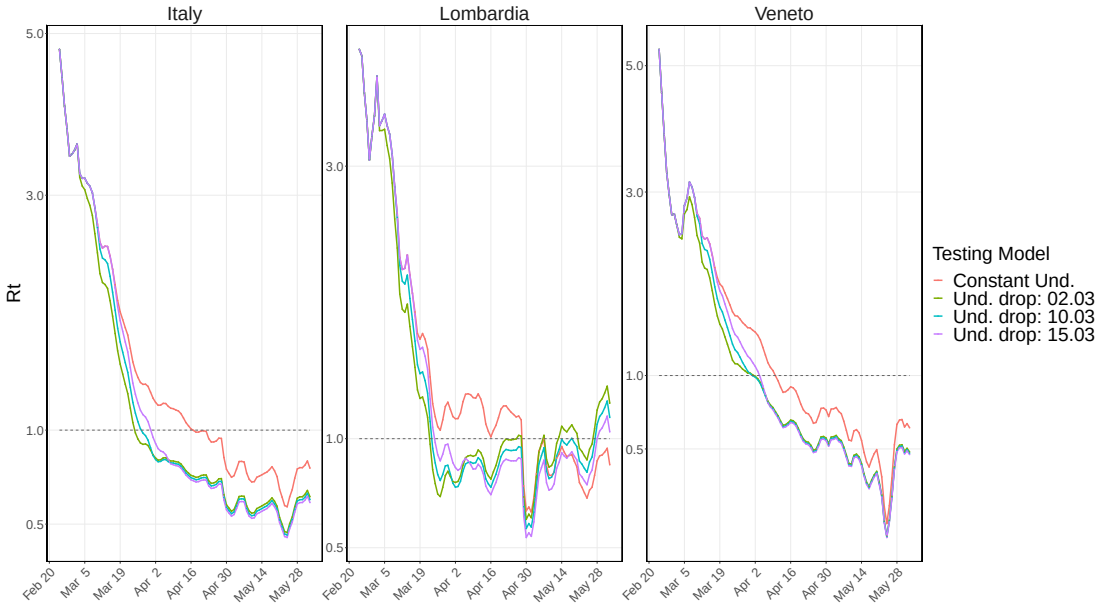

Fig S8: **Impact of testing upon R_t evolution.** R_t curves resulting from the *Testing* model in three evaluated scenarios: undetected (Und.) is decreased up to 50% starting from one week before the lockdown (green), the day of the lockdown (light blue) and one week post lockdown (lilac).

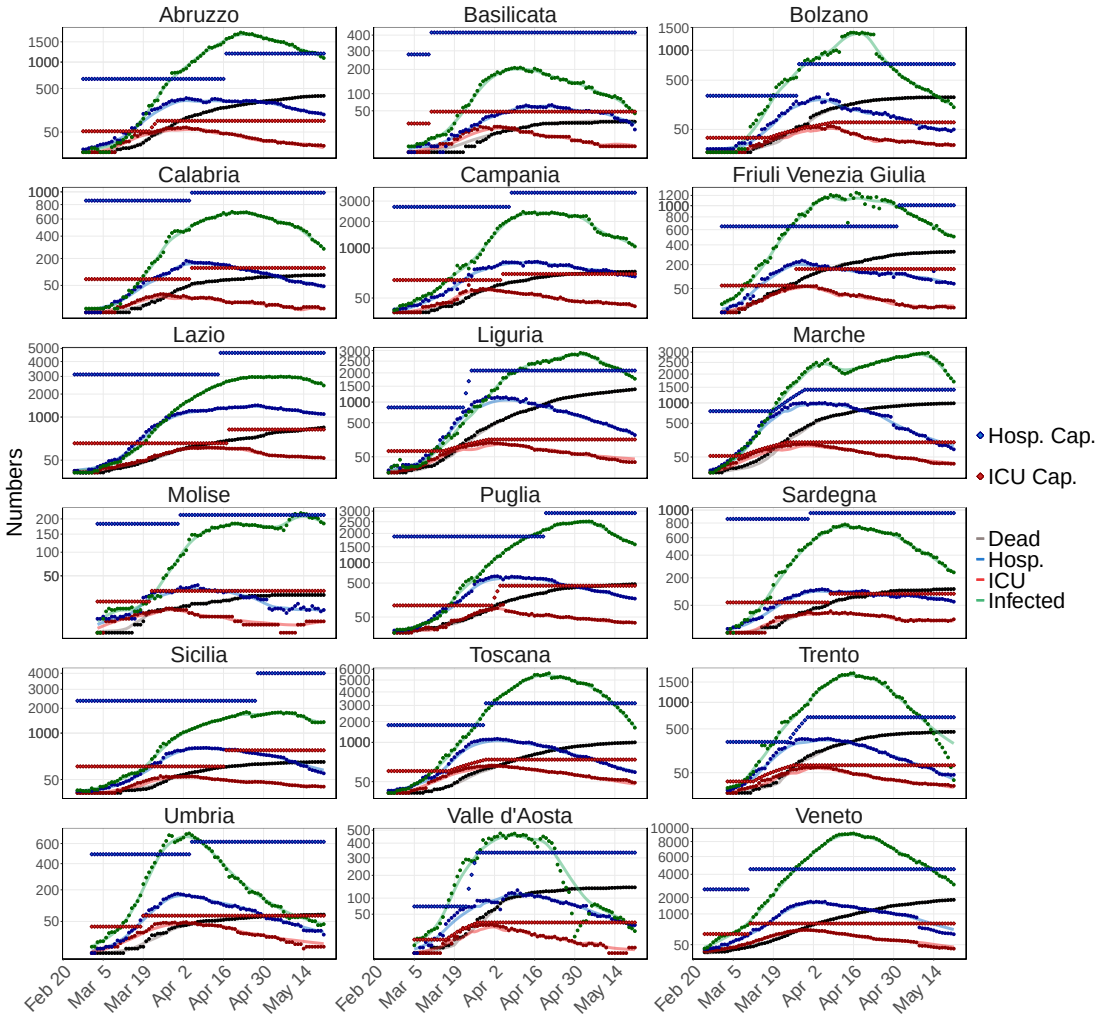

Fig S9: ***Capacity* model fit.** Fit of Infected (green), Hospitalized (light blue), ICU (red) and Death (gray) data for Italian regions (except those presented in Main text). Continuous lines: simulation results; round dots: data; diamonds: capacity limits for hospital (blue) and ICU (red).

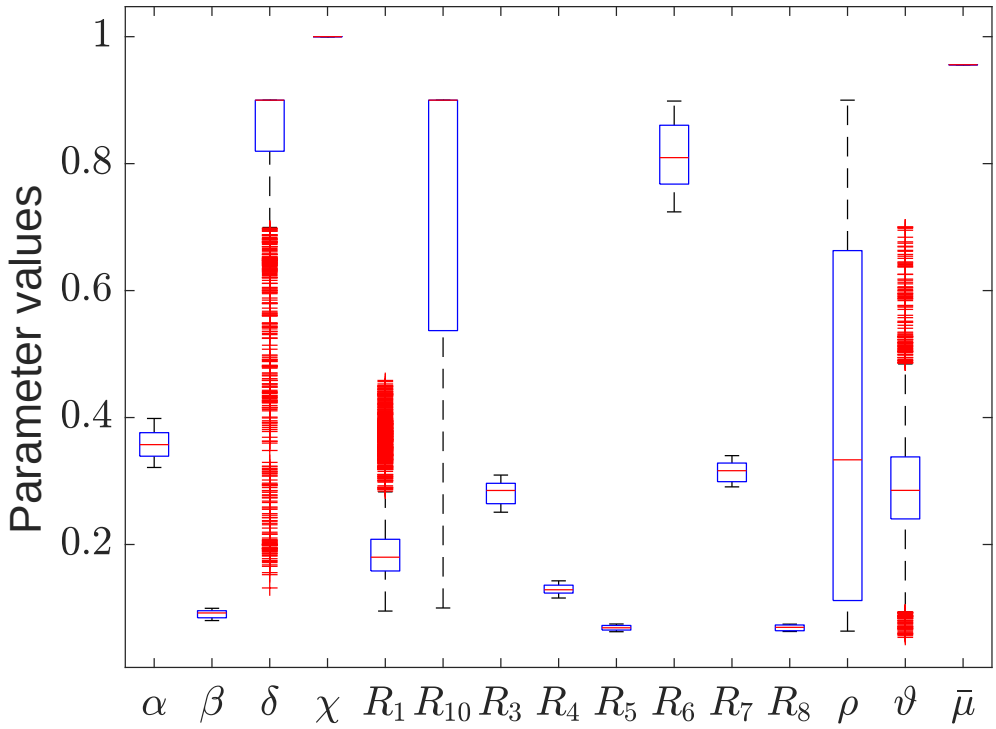

Fig S10: **Parameters' perturbation for Lombardia.** Boxplot of individual parameters resulting from the *Capacity* model fit of Lombardia data by moving window method (see Methods). $R_x, x = 3 - 8$ were estimated by fitting the first 15 data points and then perturbed ($\pm 20\%$) to generate 30 different sets of parameters. Remaining parameters were estimated in each window. Statistics are performed over 30 perturbed parameter sets (see Methods).

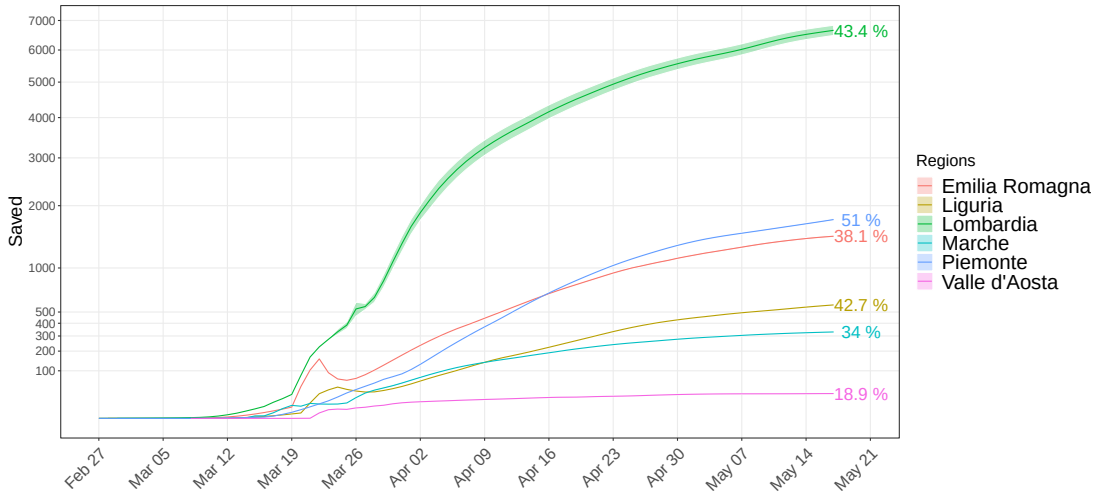

Fig S11: **Life cost of late testing with limited health care system.** We passed 30 different perturbed parameter sets of the *Capacity* model to *TestCap* model to estimate the number of infected people could have been saved by adopting early testing with increased hospital capacity. Percentages represent saved people with respect to the registered COVID-19 death. The impact is drastic in Lombardia and $\sim 43\%$ people would have been saved by adopting early testing with good infrastructure.

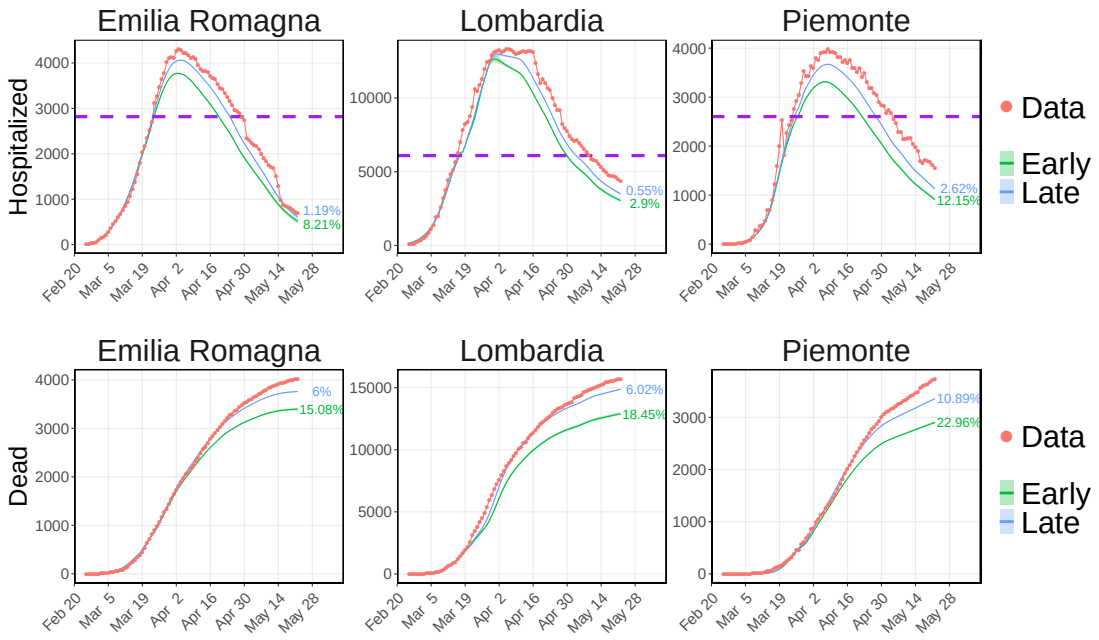

Fig S12: **Impact of 5 times more tests.** Simulation results from the *TestCap* model. The capacities are fixed at their maximum from the beginning and the undetected fraction ($\bar{\mu}$) is reduced starting one week before the lockdown (Early, green) or one week post lockdown (Late, light blue), resembling ~ 5 fold more testing than the average tests performed per day. Hospitalized is the sum of hospital and ICU patients. The reduction in Hospitalized peak and Dead number with respect to the data (red dots) has been reported as percentage. The purple horizontal line is the capacity (Hospital + ICU) before the pandemic. 5 times more testing with highest hospital beds would have decreased hospitalised at later stage of epidemic and would have saved $\sim 15\% - 23\%$ of infected people.

SECIRD models

***Reference* model equations:** Equations of the *Reference* model used to fit the data (Fig. 3A and Fig. S9) and to evaluate R_{t} (Fig. 3B and Fig. S6).

$$\frac{dS}{dt} = -R_1(t)\frac{(C_I + C_R + I_X + \beta(I_H + I_R))}{N}S \tag{1}$$

$$\frac{dE}{dt} = R_1(t)\frac{(C_I + C_R + I_X + \beta(I_H + I_R))}{N}S - R_2E \tag{2}$$

$$\frac{dC_I}{dt} = (1-\alpha)\,R_2E - R_3C_I \tag{3}$$

$$\frac{dC_R}{dt} = \alpha R_2E \;\; - R_9C_R \tag{4}$$

$$\frac{dI_H}{dt} = \frac{1-\bar{\mu}}{1-\alpha}\rho(t)R_3C_I - R_6I_H \tag{5}$$

$$\frac{dI_R}{dt} = \frac{1-\bar{\mu}}{1-\alpha}\,(1-\rho(t))\,R_3C_I - R_4I_R \tag{6}$$

$$\frac{dI_X}{dt} = \left(1 - \frac{1-\bar{\mu}}{1-\alpha}\right)R_3C_I - R_4I_X \tag{7}$$

$$\frac{dH_U}{dt} = \vartheta(t)R_6I_H - R_7H_U \tag{8}$$

$$\frac{dH_R}{dt} = (1-\vartheta(t))\,R_6I_H - R_5H_R \tag{9}$$

$$\frac{dU_D}{dt} = \delta(t)R_7H_U - R_{10}U_D \tag{10}$$

$$\frac{dU_R}{dt} = (1-\delta(t))\,R_7H_U - R_8U_R \tag{11}$$

$$\frac{dR_Z}{dt} = R_4I_R + \;\; R_5H_R + \;\; R_8U_R \tag{12}$$

$$\frac{dR_X}{dt} = R_9C_R + \;\; R_4I_X \tag{13}$$

$$\frac{dD}{dt} = R_{10}(t)U_D \tag{14}$$

***Testing* model equations:** Equations of the *Testing* model used to quantify the impact of the undetected reduction on the Hospitalised and Dead compartments (Fig. 4A,B) and on the R_{t} curve (Fig. S8). It includes an additional compartment (I_{XD}) with respect to the basic *Reference* Model which represents the additional Detected Infected.

$$
\begin{aligned}
\frac{dS}{dt} &= -R_1(t)\frac{(C_I + C_R + I_X + \beta(I_H + I_R + I_{XD}))}{N}S \\
\frac{dE}{dt} &= R_1(t)\frac{(C_I + C_R + I_X + \beta(I_H + I_R + I_{XD}))}{N}S - R_2E \\
\frac{dI_H}{dt} &= \frac{1-\bar{\mu}}{1-\alpha}\rho(t)R_3C_I - R_6I_H \\
\frac{dI_R}{dt} &= \frac{1-\bar{\mu}}{1-\alpha}\left(1-\rho(t)\right)R_3C_I - R_4I_R \\
\frac{dI_X}{dt} &= \frac{\mu'-\alpha}{1-\alpha}R_3C_I - R_4I_X \\
\frac{dI_{XD}}{dt} &= \frac{\bar{\mu}-\mu'}{1-\alpha}R_3C_I - R_4I_{XD} \\
\frac{dR_Z}{dt} &= R_4I_{XD} + R_4I_R + \ R_5H_R + \ R_8U_R
\end{aligned}
$$

***Capacity* model equations:** Equations of the *Capacity* model used to simulate the Infection dynamics with limited Hospital and ICU capacity (Fig5A, FigS9) and estimate its impact on the death toll (Fig5B). f_{Hlim} and f_{Ulim} functions are included in the equations to drive away the flux toward the dead compartment, from Infected or Hospitalised respectively, when Hospital or ICU limits are reached.

$$
\begin{aligned}
f_{Hlim} &= \left(\frac{\exp(H_U + H_R - H_{lim})^{10}}{1+\exp(H_U + H_R - H_{lim})^{10}}\right)^{10}\Bigg|_{\substack{=1 \text{ when } H_{lim}<H_R+H_U \\ =0 \text{ when } H_{lim}>H_R+H_U}} \\
f_{Ulim} &= \left(\frac{\exp(U_D + U_R - U_{lim})^{10}}{1+\exp(U_D + U_R - U_{lim})^{10}}\right)^{10}\Bigg|_{\substack{=1 \text{ when } U_{lim}<U_D+U_R \\ =0 \text{ when } U_{lim}>U_D+U_R}}
\end{aligned}
$$

$$
\begin{aligned}
\frac{dI_H}{dt} &= \frac{1-\bar{\mu}}{1-\alpha}\rho R_3C_I - R_6I_H(1-f_{Hlim}) - R_6I_Hf_{Hlim} \\
\frac{dH_U}{dt} &= \vartheta(t)R_6I_H(1-f_{Hlim}) - R_7H_U \\
\frac{dH_R}{dt} &= (1-\vartheta(t))R_6I_H(1-f_{Hlim}) - R_5H_R \\
\frac{dU_D}{dt} &= \delta(t)R_7H_U(1-f_{Ulim}) - R_{10}U_D \\
\frac{dU_R}{dt} &= (1-\delta(t))R_7H_U(1-f_{Ulim}) - R_8U_R \\
\frac{dI_D}{dt} &= R_6I_Hf_{Hlim} - R_7I_D \\
\frac{dD}{dt} &= R_{10}(t)U_D + R_7X_7f_{Ulim} + R_7I_D
\end{aligned}
$$

***TestCap* equations:** I_{XD} compartment is included in the *Capacity* Model to estimate the impact on Hospitalised and Death toll (Fig6) when the undetected infected get reduced and the hospital capacity is lifted.

$$
\begin{aligned}
\frac{dS}{dt} &= -R_1(t)\frac{(C_I + C_R + I_X + \beta(I_H + I_R + I_{XD}))}{N}S \\
\frac{dE}{dt} &= R_1(t)\frac{(C_I + C_R + I_X + \beta(I_H + I_R + I_{XD}))}{N}S - R_2E \\
\frac{dI_X}{dt} &= \frac{\mu' - \alpha}{1-\alpha}R_3C_I - R_4I_X \\
\frac{dI_{XD}}{dt} &= \frac{\bar{\mu} - \mu'}{1-\alpha}R_3C_I - R_4I_{XD} \\
\frac{dR_Z}{dt} &= R_4I_{XD} + R_4I_R + \ R_5H_R + \ R_8U_R
\end{aligned}
$$

Table S1: Parameter ranges used in the SECIR model: determination of the boundaries for literature-based parameter set was based on the interpretation of the values given in the references and is discussed in the supporting information [30, 39, 40].

Parameter	**Comments**	**Parameter ranges**	
		Min	Max
$\boldsymbol{R_1}$	Variable; fitted		
$\boldsymbol{R_2}$	$1/R2 = 5.2 - 1/R_3$; median incubation period is 5.2 days		
$\boldsymbol{R_3}$		$\frac{1}{4.2}$	$\frac{2}{5.2}$
$\boldsymbol{R_4}$		$\frac{1}{14}$	$\frac{1}{7}$
$\boldsymbol{R_5}$		$\frac{1}{16}$	$\frac{1}{5}$
$\boldsymbol{R_6}$		$\frac{1}{7}$	0.9
$\boldsymbol{R_7}$		$\frac{1}{3.5}$	1
$\boldsymbol{R_8}$		$\frac{1}{16}$	$\frac{1}{3}$
$\boldsymbol{R_9}$	$\frac{1}{R_9} = \frac{1}{R_3} + \left(0.5 \times \frac{1}{R_4}\right)$		
$\boldsymbol{R_{10}}$	Variable; fitted	$\frac{1}{10}$	0.9
$\boldsymbol{\alpha}$	fixed	0.4	0.4
$\boldsymbol{\beta}$	Assumed	0.05	1
$\boldsymbol{\delta}$	Variable; fitted	0.3	0.9
$\boldsymbol{\rho}$	Variable; fitted	0.01	0.9
$\boldsymbol{\vartheta}$	Variable; fitted	0.01	0.7
$\boldsymbol{\bar{\mu}}$	Estimated through MLE method of Bayesian framework		