



The SIR model towards the data

One year of Covid-19 pandemic in Italy case study and plausible “real” numbers

Ignazio Lazzizzera^{1,2,a}

¹ Department of Physics, University of Trento, via Sommarive 14, Povo (Trento), Italy

² Present address: Trento Institute for Fundamental Physics and Applications, INFN, via Sommarive 14, 38123 Povo, (TN), Italy

Received: 8 June 2021 / Accepted: 21 July 2021

© The Author(s) 2021

Abstract In this work, the SIR epidemiological model is reformulated so to highlight the important *effective reproduction number*, as well as to account for the *generation time*, the inverse of the *incidence rate*, and the *infectious period* (or *removal period*), the inverse of the *removal rate*. The aim is to check whether the relationships the model poses among the various observables are actually found in the data. The study case of the second through the third wave of the Covid-19 pandemic in Italy is taken. Given its scale invariance, initially the model is tested with reference to the curve of swab-confirmed infectious individuals only. It is found to match the data, if the curve of the *removed* (that is healed or deceased) individuals is assumed underestimated by a factor of about 3 together with other related curves. Contextually, the *generation time* and the *removal period*, as well as the *effective reproduction number*, are obtained fitting the SIR equations to the data; the outcomes prove to be in good agreement with those of other works. Then, using knowledge of the proportion of Covid-19 transmissions likely occurring from individuals who didn't develop symptoms, thus mainly undetected, an estimate of the *real numbers* of the epidemic is obtained, looking also in good agreement with results from other, completely different works. The line of this work is new, and the procedures, computationally really inexpensive, can be applied to any other national or regional case besides Italy's study case here.

1 Introduction

The SIR model [1–6], developed by Kermack and McKendrick [1] in 1927, is the well-known very simple model of infectious diseases that considers *three-compartments*, recalled here to state terminology and notations:

The compartment **S** of susceptible individuals;

The compartment **I** of the *infectious* (or *currently positive*) individuals, who have been infected and are capable of infecting susceptible individuals during the *infectious period*;

The compartment **R** of the *removed* individuals, who recovered from the disease or died from the disease, the former assumed to remain immune afterwards.

^a e-mail: ignazio.lazzizzera@unitn.it (corresponding author)

Births and non-epidemic-related deaths are neglected.

The cardinality of each of the compartments is indicated with the corresponding non bold letters, while N denotes the involved total population at an initial time t_0 :

$$S(t_0) + I(t_0) + R(t_0) = N. \quad (1)$$

The disease *incidence rate* β is defined so that $\beta S I$ gives the number of new infections per unit time [5]; the *removal rate* γ is defined so that γI gives the rate at which infectious individuals “deactivate” (heal or die). Typically, β is taken constant over time, which is not the general case, due to possible mutations of the disease carrier or social measures to counter the spread of the infection; also, to simplify mathematics, the *generation time* \mathbf{g} is neglected, that is the infector-infected pairing time lapse; as well as the *removal period* \mathbf{r} , which is the average time between infection and recovery or death, despite the important relation

$$\mathbf{r} = \frac{1}{\gamma} \quad (2)$$

Within the removal period \mathbf{r} , a typical infectious individual is expected to cause $\mathbf{r} \beta S$ new infections, so defining a function of time that, normalized, is called *effective reproduction number* \mathcal{R}_t (see, for instance, [7]), namely:

$$\mathcal{R}_t = \mathbf{r} \beta(t) \frac{S(t)}{N} = \frac{\beta(t)}{\gamma} \frac{S(t)}{N}. \quad (3)$$

So, the SIR equations as used here become:

$$\frac{dS}{dt}(t + \mathbf{g}) = -\gamma \mathcal{R}_t I(t), \quad (4a)$$

$$\frac{dI}{dt}(t + \mathbf{g}) = \gamma \mathcal{R}_t I(t) - \gamma I(t + \mathbf{g} - \mathbf{r}), \quad (4b)$$

$$\frac{dR}{dt}(t + \mathbf{r}) = \gamma I(t). \quad (4c)$$

Of course, they imply that the sum $S + I + R$ is conserved, so that $S(t) + I(t) + R(t) = N$ at any time t .

2 Outlines of the work

The purpose of this work is to check whether the relations established in Eq. 4 are actually found in the data or, at least, whether “corrections”, accounting for incomplete data or systematic errors, may or should be introduced, with the implication that consequently the relationships are satisfied. Crucial is the fact that the model is scale-invariant, thus allowing to conveniently choose as a reference one sub-compartmental curve whose real data can be considered reliable, such as the swab-confirmed infectious individuals curve. This choice is done indeed here: swab-confirmed infectious are mostly individuals who have developed symptoms and are actually found to cover a nearly constant fraction of all the infectious people, given the circumstances that symptomatic and asymptomatic individuals roughly are, respectively, fractions of the age groups of over sixty and younger people ([8–11]).

The case of the second through the third pandemic wave of Covid-19 in Italy is studied. First, it will be shown that the relation established through Eq. 4c holds true if $R(t)$ is scaled by a factor that is obtained, together with the *removal period* \mathbf{r} , by a least-square procedure

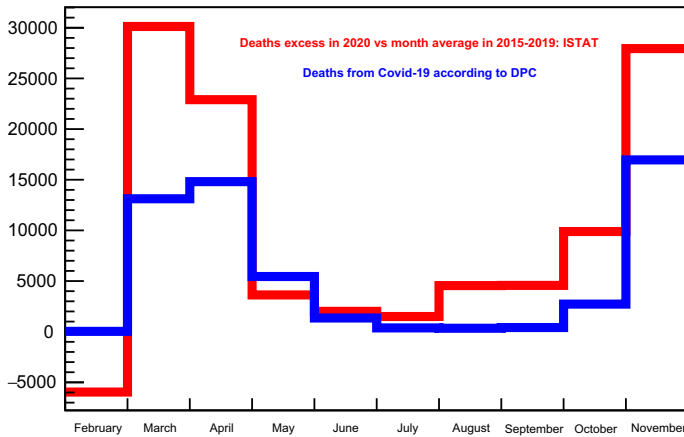


Fig. 1 Monthly deaths from Covid-19 in Italy during 2020, according Italy's Department of Protezione Civile (blue histogram); monthly excess of deaths the same year compared to the averages over the previous five years, according Italy's ISTAT (Istituto Nazionale di Statistica)

of matching over data. Arguments will be given for how a scaling-up over the official data should be due indeed.

Once r , and thus γ , is obtained, the *effective reproduction number* is evaluated through Eq. 4b, reliably, despite using the swab-confirmed infection cases only, for that equation is scale-invariant on its own.

The transition to *real* numbers is finally done, correcting the swab-confirmed infectious cases for the proportion factor of transmissions that likely occur from asymptomatic subjects. The results are compared with those obtained at the MRC Centre for Global Infectious Disease Analysis, Imperial College London (ICL, [12]), where a completely independent approach is used.

3 The data set

The data set is from Italy's Department of Protezione Civile [13], lasting from 1 June 2020 to 31 May 2021. Since every weekend there was a postponement in cases recording to a few days later, according to common practice the data are smoothed via a multi-day moving average; the choice is 9 days, to systematically include a couple of days after each weekend.

4 The swab-confirmed infectious towards the daily removed

Verifying that the relationship given by Eq. 4c is indeed found in the data is not so trivial. For example, there is evidence that the monthly deaths from Covid-19 in 2020, as given by Italy's Department of Protezione Civile, are largely underestimated: this is shown by an ISTAT study on the monthly excess of deaths in 2020, compared to the corresponding averages over the previous five years (see [14] and [15]). ISTAT is Italy's Istituto Nazionale di Statistica. The matter is illustrated in Fig. 1. In addition to this, it is to be expected that $R(t)$ does not include most of the cases that had an asymptomatic or mild course. Also, asymptomatic infected people are probably not reported among the infectious, whereas by

Table 1 k-factor and \mathbf{r}

	k_{rel}	\mathbf{r}	χ^2	Ndf	red. χ^2
Likelihood on raw data	2.99 ± 0.82	10.31 ± 2.74	93185	360	266
Likelihood on smoothed	3.14 ± 0.82	10.32 ± 2.68	32823	360	91
Gaussian fit		9.78 ± 2.28	1.15	32	0.04
Skew sigm. derivative		10.14 ± 2.37	2.48	66	0.04

far most of the reported infectious are those who had swabs confirmation, whose number will be called I_{sc} . Let's indicate with \check{R} the curve of the actually registered healed and deaths: it is found that its derivative, the daily variation, shifted forward in time, is indeed proportional to $I_{\text{sc}}(t)$. To methodically verify Eq. 4b, the correction factor k_{rel} is introduced so as to give maximum generality to a least-square search over the positive definite form

$$\chi^2(k_{\text{rel}}, \mathbf{r}) = \sum_b \left[I_{\text{sc}}(b) - k_{\text{rel}} \mathbf{r} \frac{d\check{R}}{dt}(b + \mathbf{r}) \right]^2, \quad (5)$$

with varying k_{rel} and the *removal period* \mathbf{r} . It is worth remarking the notation k_{rel} , intended to emphasize that any correction on $\check{R}(t)$, possibly *required* by the SIR model at this stage, is relative to the swab-confirmed infectious population only. The sum is over the days of the pandemic period considered, with the choice of weighing equally all daily data. In principle two functions of the kind 5 should be used, one for deaths and one for recoveries; however, deaths are well known to be only at most some 4–5% of the whole removed compartment, so that a probably negligible error is made by simplifying as in Eq. 5 because of the huge statistical and systematic uncertainties in the data. The minimization is performed using a C++ object of the class *Minimizer* of the CERN package ROOT, typically used by high energy physicist in their data analysis ([16, 17]): its statistical methodology is described in [18]. Since the surface defined from the data through Eq. 5 is rather rough, the minimization algorithm is run 150,000 times to maximize the chance of hitting an optimal minimum: the initial values of k_{rel} and \mathbf{r} are drawn at random in the intervals [1.0, 5.0] and [5.0, 18.0] respectively. Execution on raw and smoothed data takes about one minute time altogether. The final issue for k_{rel} and \mathbf{r} and their uncertainties δk_{rel} and $\delta \mathbf{r}$ are taken as the mean and the standard deviation of the distributions of the respective outcomes at each iterated minimization, weighted with the normalized inverse of the χ^2 .

The results are shown on the first and the second lines of Table 1, for the raw and the smoothed data respectively.

Since the value of the *removal period* \mathbf{r} is critical in determining k_{rel} , it is sought from the data in two further independent ways, as explained in the next two subsections.

4.1 The removal period from a Gaussian fit

At any new “wave” of epidemic, the rise in number of the infectious individuals follows with good approximation a sigmoidal shape, i.e. it is roughly exponential at the very beginning, up to an inflection point, after which it bends towards a plateau; consequently, its daily variation (the time derivative) exhibits a maximum at the inflection point, around which it is approximately Gaussian. If Eq. 4c correctly described the data, an analogous shape should be had in the second derivative of the *removal curve*. Very remarkably, this is in fact the case,

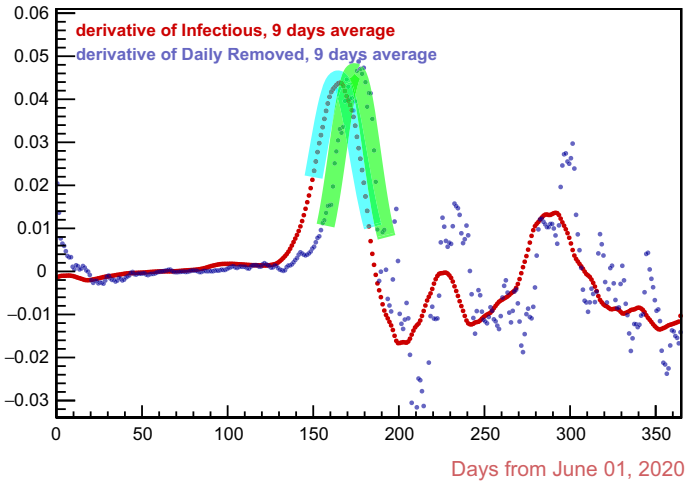


Fig. 2 The two fitting Gaussian functions in cyan and green

as shown in Fig. 2, where fitting Gaussians are plotted over the first derivative of the *infectious curve* and the second derivative of the *removal curve*: the distance in time between the vertexes of the two Gaussians gives a new independent *measurement of the removal period r*, reported in the third line of Table 1 with associated uncertainty and fit χ^2 . The uncertainty is the sum in quadrature of the uncertainty on the position in time of the vertexes of the two fitting Gaussians; the χ^2 , with its reduced, is their worse. The fit algorithm is from the already mentioned ROOT package (CERN, [19]).

4.2 The removal period from the “asymmetric sigmoid derivative” fit

Given the almost sigmoidal initial growth of an epidemic wave, as already mentioned in the last subsection, an alternative fit function turns to be an asymmetric modification of the derivative of a sigmoid, which will be called *skew sigmoid derivative*, namely:

$$\begin{cases} s(x; \mu, \sigma, \epsilon) = \frac{1}{\epsilon + (1 - \epsilon)e^{-(x-\mu)/\sigma}}, & \text{with } 0 \leq \epsilon < 1, \\ \mathcal{A}(x; \mu, \sigma, \epsilon) = A s(x; \mu, \sigma, \epsilon) [2 - s(x; \mu, \sigma, \epsilon)]. \end{cases} \quad (6)$$

This function has absolute maximum in $x = \mu$, with $\mathcal{A}(\mu; \mu, \sigma, \epsilon) = A$. It is plotted in Fig. 3 for $\mu = 1$, $\sigma = 1$ and $A = 1$, and various values of the *skewness* parameter ϵ : for $\epsilon = 0.5$ one has the derivative of a very sigmoid. The fits of the skew sigmoid derivative to the first derivative of the swab-confirmed *infectious curve* and to the second derivative of the *removal curve*, respectively, are shown in Fig. 4: again, the distance in time between the vertexes of the fitting functions gives a new *measurement of the removal period r*, reported in the fourth line of Table 1.

4.3 The removal curve corrected relatively to the swab-confirmed infectious only

From Table 1 the *removal period r* is assumed to be 10 ± 2 days, bearing in mind that the data have just one day resolution; also, comparing the χ^2 on the first and second lines of the table, the correction factor k_{rel} is taken equal to 3.14 ± 0.82 . So, we have the curve of the

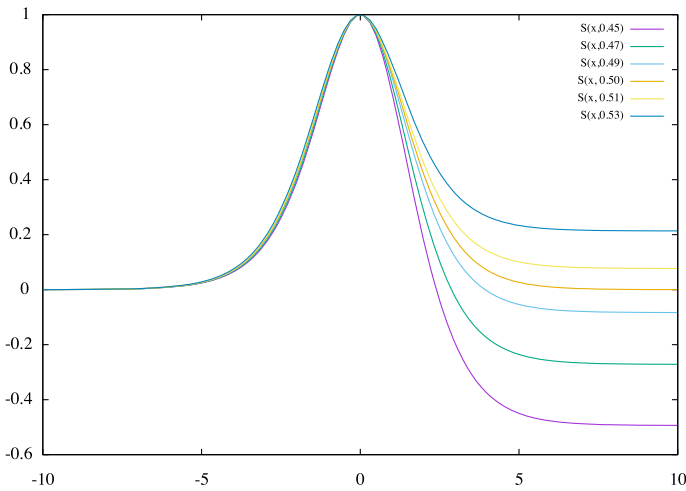


Fig. 3 The skew sigmoid derivative for $\mu = 1.$, $\sigma = 1.$, $A = 1.$; $\epsilon = 0.45, 0.47, 0.49, 0.50, 0.53$

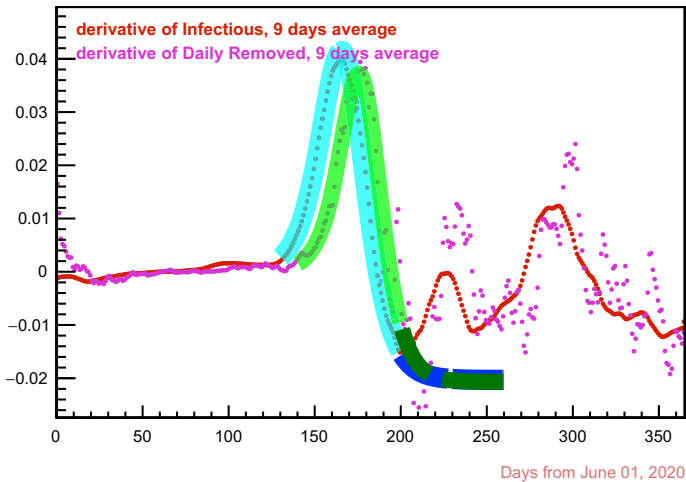


Fig. 4 The two fitting “asymmetric sigmoid derivative” functions in cyan and green

removed individuals, corrected relatively to the swab-confirmed infectious only, given by:

$$R_{rel}(t) = k_{rel} \check{R}(t) \quad \ni \quad \frac{dR_{rel}}{dt}(t + \mathbf{r}) = \gamma I_{sc}(t). \tag{7}$$

Figure 5 does illustrate this: the cyan error bars are generated by the propagation of three times the ± 0.82 uncertainty over k_{rel} .

5 Getting the generation time and the effective reproduction number

There are several algorithms to estimate the *effective reproduction number* from the data: a simplified one is given in [20], where also an extensive bibliography on the subject can be

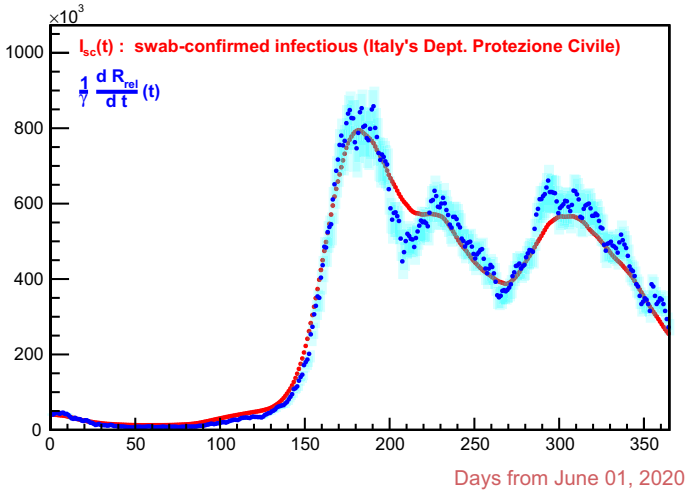


Fig. 5 Relation 7 in the data. The derivative of $R_{rel}(t)$ (blue marks) is time-shifted by the removal period \mathbf{r} . The cyan error bars are given by propagation from three times ± 0.82 uncertainty over k_{rel}

found. The simplest yet effective estimate that very directly interprets the meaning of the function (see, for instance, [21]) is given by

$$\mathcal{R}_t = \frac{I(t + \mathbf{g})}{I(t)}. \tag{8}$$

As far as the SIR model is concerned, from Eq. 4b one has

$$\mathcal{R}_t = \frac{\mathbf{r}}{I(t)} \frac{I(t + \mathbf{g} + 1) - I(t + \mathbf{g} - 1)}{2} + \frac{I(t + \mathbf{g} - \mathbf{r})}{I(t)}. \tag{9}$$

So, the derivative is implemented by the symmetric difference quotient, to have the cancellation of the first-order error in the numerical discretization.

While only the *generation time* \mathbf{g} appears in Eq. 8, both \mathbf{g} and the *removal period* \mathbf{r} are present in Eq. 9; consequently, the validity test of the SIR model through the *effective reproduction number* \mathcal{R}_t it manages to provide, is to be considered quite stringent.

In the previous section, the *removal period* \mathbf{r} was obtained from the data using the SIR model; the question is how to get the *generation time* as well.

In our conventions, $I(t)$ denotes the total of all the infectious people, swab-confirmed or not. If t_M is a day when $I(t)$ presents a maximum, then correspondingly, but \mathbf{g} days earlier, i.e. at day $t_M - \mathbf{g}$, the *effective reproduction number* \mathcal{R}_t should be equal to 1, because an increase in the number of the people becoming infectious requires $\mathcal{R}_t > 1$ and a decrease requires $\mathcal{R}_t < 1$. Of course, every variation of \mathcal{R}_t has impact on $I(t)$ with a delay of \mathbf{g} days, so also for $I_{sc}(t)$, assuming this to be proportional to $I(t)$. With \mathbf{r} fixed at 9.71 ± 2 days, as set out in the previous section, let's say t_g a day when, for any given choice of \mathbf{g} , \mathcal{R}_t is equal to 1: in general, checking over the data, it doesn't happen that the nearest next day t_M , on which $I_{sc}(t)$ has a maximum, is such that $t_M - t_g = \mathbf{g}$, as it should; indeed it happens only for a specific choice of \mathbf{g} , namely, for the case being studied, with $\mathbf{g} = 6$, an integer value just in view of the one-day resolution of the data. A convenient double check is done on the maximum of $I_{sc}(t)$ falling on 2 December 2020 (see Fig. 5). Very remarkable is the fact that the height of the peaks of \mathcal{R}_t does depend on the value one wants to give to \mathbf{g} , the same way

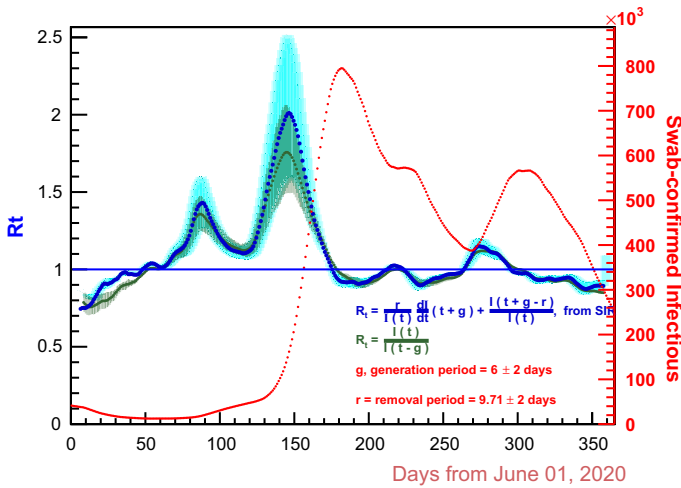


Fig. 6 \mathcal{R}_t according to the SIR model, compared with the used simplest formula, Eq. 8. In cyan are the error bars for Eq. 9; the other are in black. $g = 6 \pm 2$; $r = 10 \pm 2$

as the days when \mathcal{R}_t is equal to 1 do: so, all of these things are bounded by the SIR model, a fact that must be considered truly important in evaluating the validity of the model.

The estimate $g = 6$ days is in total agreement with the average 6.7 ± 1.9 days given for Italy in Ref. [22]: this is a success of Eq. 9 that strengthens the agreement, within the uncertainties, of the resulting \mathcal{R}_t with that from other algorithms, as those reported in Ref. [20], with references therein. Figure 6 shows this SIR generated \mathcal{R}_t , together with the one from Eq. 8; for either, error bars corresponding to a ± 2 days uncertainty on both g and r are also shown.

6 The “corrected” cumulative and daily-new infections relatively to the swab-confirmed infectious people

To avoid confusion, it is worth remarking that *infections at day t* is meant as the cumulative number of infections up to and including that day, while the number of *infectious people* at some day t refers to those people who were infected possibly earlier and are still able to transmit infection at that day. Thus, the (daily new) *infections curve* is different from the *infectious curve*.

Since N in Eq. 1 is conserved, Eq. 4a can be written as

$$\frac{d(N - S)}{dt}(t + g) = \gamma \mathcal{R}_t I(t), \tag{10}$$

expressing the daily new infections, gross of *removed* people (while the *infectious* numbers are *net of removals*). Indeed,

$$\mathcal{T}(t) = N - S(t) = I(t) + R(t) \tag{11}$$

is nothing but the total cases of infections at time t .

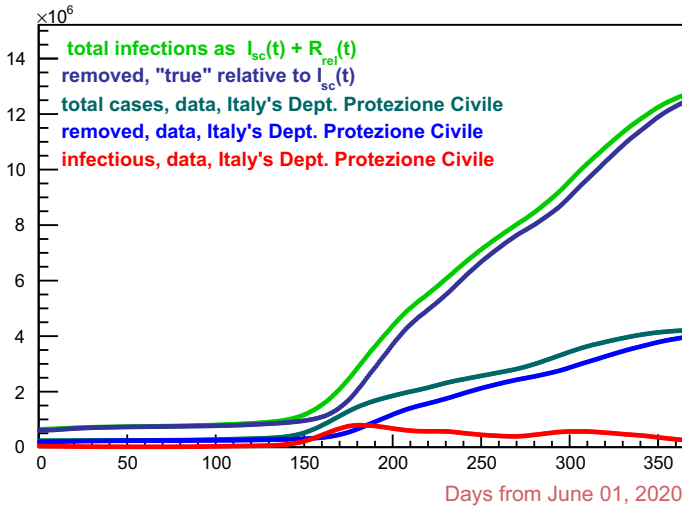


Fig. 7 Italy, Covid-19 second through the third waves: estimates of *removal* (dark blue curve) and total (green curve) cases as from correction relative to the swab-confirmed infectious (red) curve. Also shown data in dark green (total cases) and blue (*removals*)

For what has been done so far, Eq. 10 must be replaced by

$$\frac{d \mathcal{T}_{rel}(t)}{dt} = \gamma \mathcal{R}_t I_{sc}(t), \tag{12a}$$

$$\mathcal{T}_{rel}(t) = I_{sc}(t) + R_{rel}(t), \tag{12b}$$

giving the *corrected* cumulative number of infections relatively to the swab-confirmed infectious people only. Figure 7 illustrates Eq. 12.

7 Infections from asymptomatic and symptomatic infectious people: estimate of the “real” numbers

There are several studies on the relevance of SARS-CoV-2 transmission from asymptomatic people, like [8,23–25] and references therein. Quite recent and complete is ref. [25], where a decision analytical model is used to assess the proportion of SARS-CoV-2 transmissions in the community likely occurring from subjects who did not develop any symptom. In that work data from a meta-analysis were used to set the *generation time* at a median of 5 days and *infectious period* at 10 days, in good agreement, respectively, with the 6 and 10 days stated in the present work. The reported conclusion is that, across a range of plausible scenarios, a 59% of infection transmission occurs from persons without symptoms: no clear uncertainty is given, but the statement that the figure should be *at least 50%*, suggests an uncertainty of $\pm 10\%$. Also it is stated that the infected individuals who never develop symptoms are 75% as infectious as those who do develop symptoms.

Let’s call $f_{(asy)}$ the percentage fraction of the asymptomatic infectious subjects over all the infectious people and $i_{(asy)}$ their relative infectiousness, that is the percentage fraction of the infectiousness of those who had developed symptoms: then, according to the best

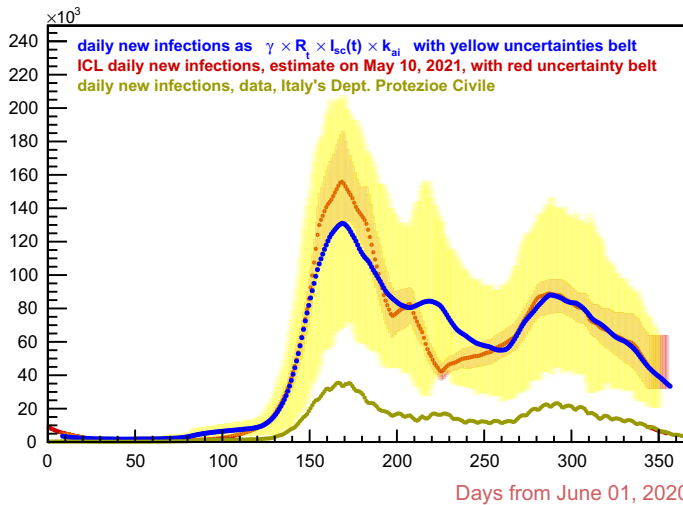


Fig. 8 Italy, Covid-19 second through the third waves. Estimates of daily new infections in this work are in blue and uncertainty belt in yellow; ICL is in dark red

available figures, it would be $f_{(asy)} = 59 \pm 10\%$ and $i_{(asy)} = 75\%$, the latter with a trial, very conservative, uncertainty of $\pm 20\%$.

Denoting by $I_{(asy)}(t)$ the number of the asymptomatic infectious individuals and recalling that $I_{sc}(t)$ indicates the number of the swab-confirmed infectious subjects, it should be

$$I(t) - I_{(sc)}(t) = I_{(asy)}(t) = \frac{f_{(asy)}}{100} \frac{i_{(asy)}}{100} I(t),$$

whence

$$I(t) \cong \left(1 + \frac{f_{(asy)}}{100} \frac{i_{(asy)}}{100} \right) I_{(sc)}(t) =: k_{(ai)} I_{(sc)}(t), \tag{13a}$$

$$R(t) = k_{(ai)} R_{rel}(t), \tag{13b}$$

$$k_{(ai)} =: 1 + \frac{f_{(asy)}}{100} \frac{i_{(asy)}}{100} \approx 1.44 \pm 0.22. \tag{13c}$$

Then, in view of Eq. 11, Eq. 10 becomes

$$\frac{d\mathcal{T}}{dt}(t + \mathbf{g}) = \gamma \mathcal{R}_t k_{ai} I_{sc}(t), \tag{14}$$

with $\mathcal{T}(t)$ the “real” cumulative number of infections at day t , while its derivative represents the “real” daily new infections.

Figure 8 shows the daily new infections curve, compared with the Imperial College’s (ICL) model estimate, as published in [12]. The model in question is a stochastic SEIR variant that adopts multiple infectious states, which in turn reflect different COVID-19 severities. It uses an estimate of the *infectious fatality rate* (IFR), assuming that the number of confirmed deaths from Covid-19 is equal to the *real* Covid-19 deaths number; it also uses an estimate of the *effective reproduction number*, based on the changes of the virus transmission rate caused by the average mobility trends.

So, the ICL model’s approach is totally different from the one followed in the present work; nevertheless the respective “real” daily new infections estimates appear to be in quite

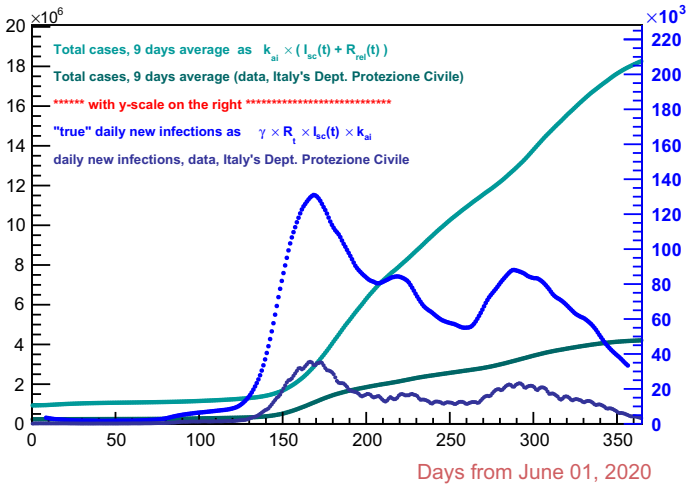


Fig. 9 This work estimates of the “real” total cases (light green) and correspondent data form Italy’s Department of Protezione Civile (dark green) with scale on the left; this work “real” daily new infections (blue) and corresponding official data (dark blue) with scale on the right

good agreement, except on a time interval around 1 January 2021 (day 220 in plots of this paper), where the ICL curve shows a deep local minimum instead of a local maximum as in the data of Italy’s Department of Protezione Civile. The uncertainty belt of the ICL estimates is surprisingly narrow. In Fig. 9 the present work’s estimates of the “real” total cases of infections are shown, together with the estimated “real” daily new infections, the latter with their own scale on the right; also, the data as from Italy’s Department of Protezione Civile are plotted.

Incidentally, the ripple visible in Figs. 8 and 9 on the data from Italy’s Department of Civil Protection has the typical 7-day periodicity that arises from the weekend reduced data recording.

8 Conclusions

Taking as case study the second to the third waves of SARS-CoV-2 in Italy, the SIR model is confronted with data, after reformulating its equations by the explicit introduction of the important *effective reproduction number* \mathcal{R}_t , as well as the *generation time* and the *infectious period*, usually, erroneously, neglected. The relationships it sets among the main observables are actually found in the data, in particular between the curve of the swab-confirmed infectious individuals and the curve of the *removed* (healed or deceased) subjects. Indeed, taking advantage of its scale invariance and choosing the curve of the swab-confirmed infectious people as a reference, the model suggests a correction on the number of *removed* individuals for just a factor which would take into account: (a) infected people who have not developed relevant symptoms and, therefore, were not detected; (b) deaths erroneously not attributed to Covid-19. *Generation time*, *infectious period* and *effective reproduction number* have been sought from the data through the model. At the very end, the curve of the swab-confirmed infectious individuals has been completed for the proportion of infection transmissions likely occurred from individuals with no symptoms, using figures published

in important works ([8, 23–25]). Thus, an estimate of the *real numbers* of the pandemic in Italy is obtained for the considered period of time. All the results are in good agreement with those of other studies, in particular of the ICL group ([12]), whose approach is totally different from the present. The vision on and use of the SIR model of this work are new; the C++ code, computationally really inexpensive and available under request to the author, can be applied to any other national or regional case besides Italy's study case here.

Funding Open access funding provided by Università degli Studi di Trento within the CRUI-CARE Agreement.

Data Availability Statement This manuscript has associated data in a data repository. [Authors comment: Original data are those from Italy's Department of "Protezione Civile", available at a public repository according to reference [13] in the paper. The resulting estimation data, as shown in the paper, are available from the author under request to him.]

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. W.O. Kermack, A.G. McKendrick, Contribution to the mathematical theory of epidemics. Proc. R. Soc. A **115**, 700–721 (1927). <https://doi.org/10.1098/rspa.1927.0118>
2. J. Murray, *Mathematical Biology* (Springer-Verlag, Berlin, 1993)
3. D. Daley, J. Gani, *Epidemic Modelling* (Cambridge University Press, Cambridge, 1999)
4. F. Brauer, Mathematical epidemiology: past, present, and future. Infect. Dis. Model. **2**, 113–127 (2017). <https://doi.org/10.1016/j.idm.2017.02.001>
5. M. Martcheva, *An Introduction to Mathematical Epidemiology*, vol. 61 (Springer, New York, 2015)
6. F. Brauer, C. Castillo-Chavez, Z. Feng, *Mathematical Models in Epidemiology* (Springer, New York, 2019)
7. H.W. Hethcote, The mathematics of infectious diseases. SIAM Rev. **42**(4), 599–653 (2000). <https://doi.org/10.1137/S0036144500371907>
8. B. Snider, B. Patel, E. McBean, Asymptomatic cases, the hidden challenge in predicting COVID-19 caseload increases. Infect. Dis. Rep. **13**, 340–347 (2021). <https://doi.org/10.3390/idr13020033>
9. P. Poletti, M. Tirani, D. Cereda et al., Association of age with likelihood of developing symptoms and critical disease among close contacts exposed to patients with confirmed SARS-CoV-2 infection in Italy. JAMA Netw Open **4**(3), e211085 (2021). <https://doi.org/10.1001/jamanetworkopen.2021.1085>
10. B. Unim, L. Palmieri, C. Lo Noce, S. Brusaferrò, G. Onder, Prevalence of COVID-19-related symptoms by age group. Aging Clin. Exp. Res. **33**(4), 1145–1147 (2021). <https://doi.org/10.1007/s40520-021-01809-y>
11. M. Monod, A. Blenkinsop, X. Xi et al., Age groups that sustain resurging COVID-19 epidemics in the United States. Science (2021). <https://doi.org/10.1126/science.abe8372>
12. <https://ourworldindata.org/covid-models#imperial-college-london-icl>
13. <https://github.com/pcm-dpc/COVID-19>
14. <https://public.tableau.com/profile/istat.istituto.nazionale.di.statistica>
15. G. Bonifazi, L. Lista, D. Menasce, M. Mezzetto, A. Oliva, D. Pedrini, R. Spighi, A. Zoccoli, A statistical analysis of death rates in Italy for the years 2015–2020 and a comparison with the casualties reported from the COVID-19 pandemic. Infect. Dis. Rep. **13**(2), 285–301 (2021). <https://doi.org/10.3390/idr13020030>
16. <https://root.cern/root/html/doc/guides/users-guide/ROOTUsersGuide.html>
17. https://root.cern.ch/doc/v614/classROOT_1_1Math_1_1Minimizer.html
18. Jacoboni, C.W.T. Eadie, D. Dryard, F.E. James, M. Roos, B. Sadoulet, Statistical methods in experimental physics. Nuov Cim. A **40**, 235 (1977). <https://doi.org/10.1007/BF02776791>

19. <https://root.cern.ch/root/html534/guides/users-guide/FittingHistograms.html>
20. G. Bonifazi, L. Lista, D. Menasce, M. Mezzetto, D. Pedrini, R. Spighi, A. Zoccoli, A simplified estimate of the effective reproduction number R_t using its relation with the doubling time and application to Italian COVID-19 data. *Eur. Phys. J. Plus* **136**(4), 386 (2021). <https://doi.org/10.1140/epjp/s13360-021-01339-6>
21. Robert Koch Institut, "Erläuterung der Schätzung der zeitlich variierenden Reproduktionszahl" https://www.rki.de/DE/Content/InfAZ/N/Neuartigesc_ornavirus/Projekte_RKI,R-Wert-Erlaeuterung.pdf (2020)
22. D. Cereda et al., The early phase of the COVID-19 outbreak in Lombardy, Italy. [arXiv:2003.09320](https://arxiv.org/abs/2003.09320) (2020)
23. D.P. Oran, E.J. Topol, Prevalence of Asymptomatic SARS-CoV-2 Infection. *Ann. Intern. Med.* (2021). <https://doi.org/10.7326/M20-3012>
24. M. Alene, L. Yismaw, M.A. Assemie, D.B. Ketema, B. Mengist, B. Kassie, T.Y. Birhan, Magnitude of asymptomatic COVID-19 cases throughout the course of infection: A systematic review and meta-analysis. *PLoS ONE* **16**(3), e0249090 (2021). <https://doi.org/10.1371/journal.pone.0249090>
25. M.A. Johansson, T.M. Quandelacy, S. Kada et al., SARS-CoV-2 transmission from people without COVID-19 symptoms. *JAMA Netw. Open* **4**(1), e2035057 (2021). <https://doi.org/10.1001/jamanetworkopen.2020.35057>
26. J.A.P. Heesterbeek, A brief history of R_0 and a recipe for its calculation. <https://doi.org/10.1023/A:1016599411804>
27. L. Edelstein-Keshet, *Mathematical Models in Biology* (Society for Industrial and Applied Mathematics, Philadelphia, 1988)
28. <https://opendatadpc.maps.arcgis.com/apps/dashboards/b0c68bce2cce478eaac82fe38d4138b1>