LETTER TO THE EDITOR

Unreliable predictions about COVID-19 infections and hospitalizations make people worry: The case of Italy

In Italy, we are currently facing the so-called fourth wave of the COVID-19 epidemic. But is this one a true wave? On TV shows and newspapers, so-called experts are widely discussing what we will face in the very near future. Sometimes it seems that everyone is playing Hide and Seek. Let us introduce some history of the game.

Hide and Seek was first described by a Greek writer named Julius Pollux in the 2nd century BCE. Pollux was a Greek scholar and rhetorician who rose to success in Athens and Egypt. Getting closer to our time in Elizabethan England, the linked text mentions the game and *describes children using blindfolds*. And around the same time, *hide-and-seek makes a brief appearance in Shakespeare's* Love's Labors Lost, when he says, "All hid, All hid, an old infant play." Today we find distinguished scholars, experts in numbers, and their reading, playing the same game on the media.

Unfortunately, as also discussed in Martinez et al.,[1] the pandemic that we are facing is not a game. Members of official committees deputed to several forms of control of the pandemic situation, who oversees monitoring and proposing tools of interventions to mitigate the expansion of the epidemic, contribute to the general confusion and the unclear discussion.

In July, official declarations said that we would observe 30 000 or 40 000 cases per day in mid-August. Would you like to know the size of the incidence peak? The incidence peak of daily cases has been around 7800 cases at the end of August 2021, with an approximated range of 4000–8000 daily contagions in August (Figure 1). Furthermore, during the last 7 weeks (12.7–29.8), the percent variation of the incidence has decreased from 103% to 3.8%, through 85%, 27%, 8.3%, 7.9%, and −0.8%, respectively (Figure 2), with the ratio of cases to swabs relatively stationary and around 2.97% in August (Figure 3).

How is it possible that official predictions are so different from observed numbers? The methodology used has not been disclosed. Still, several scientific sources report that those numbers were obtained considering the patterns shown by the UK epidemic curves, that is, by similarity. Other experts also predicted pretty much the same number of daily cases, based on exponential modeling (see the blue line in Figure 1), that is, as the epidemic will never stop. Although several models have been proposed and published to analyze the Italian epidemic,[2-4] none has been considered by those who have to manage the evolution of the epidemic, to the best of our knowledge.

As a result, people are confused, do not completely trust official communications anymore, and doubt the reliability of predictions, scenarios, and projections that determine daily-life restrictions. Of course, it is not a single event to mine people's
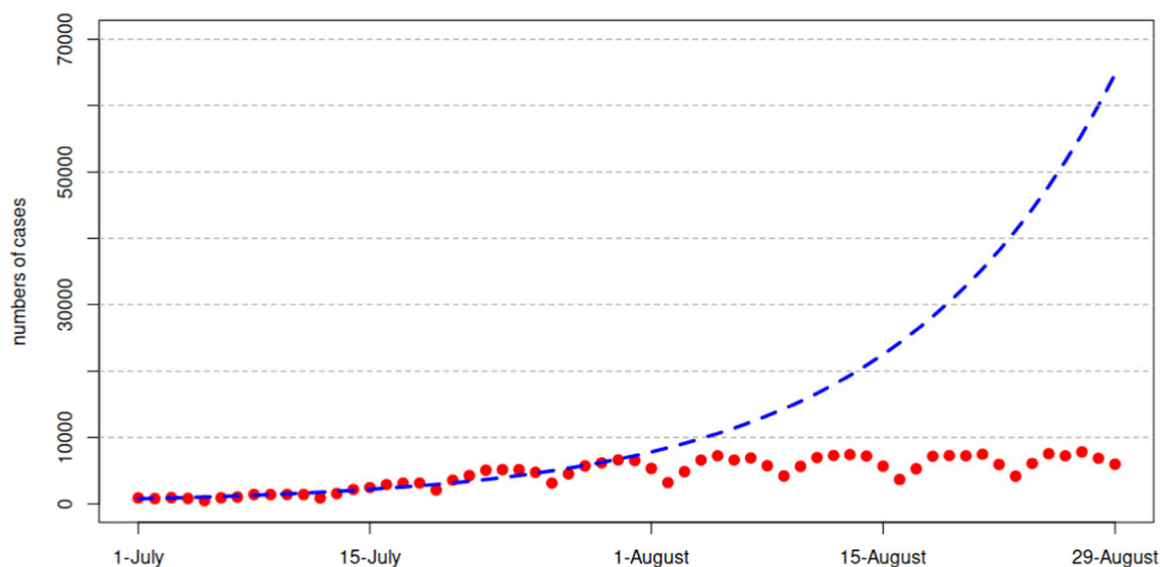

FIGURE 1 COVID-19 in Italy: Daily incidence cases in July and August 2021 (red points) versus the Poisson exponential curve estimated using the training cases observed from 1.7–31.7 (blue-dashed line)
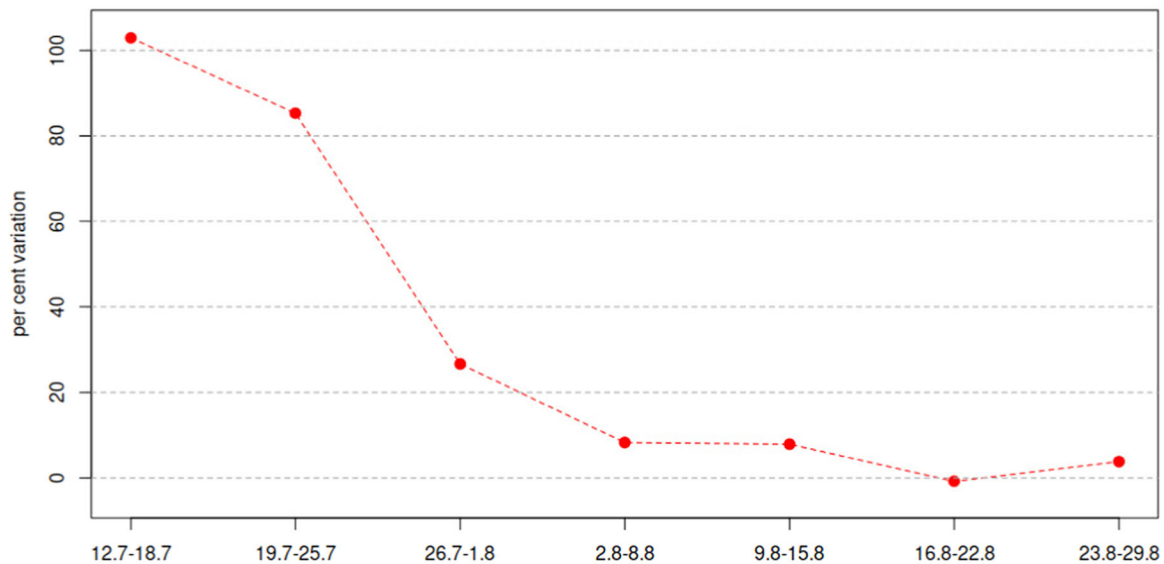
FIGURE 2 COVID-19 in Italy: Percent variation of the weekly incidence in the last 7 weeks
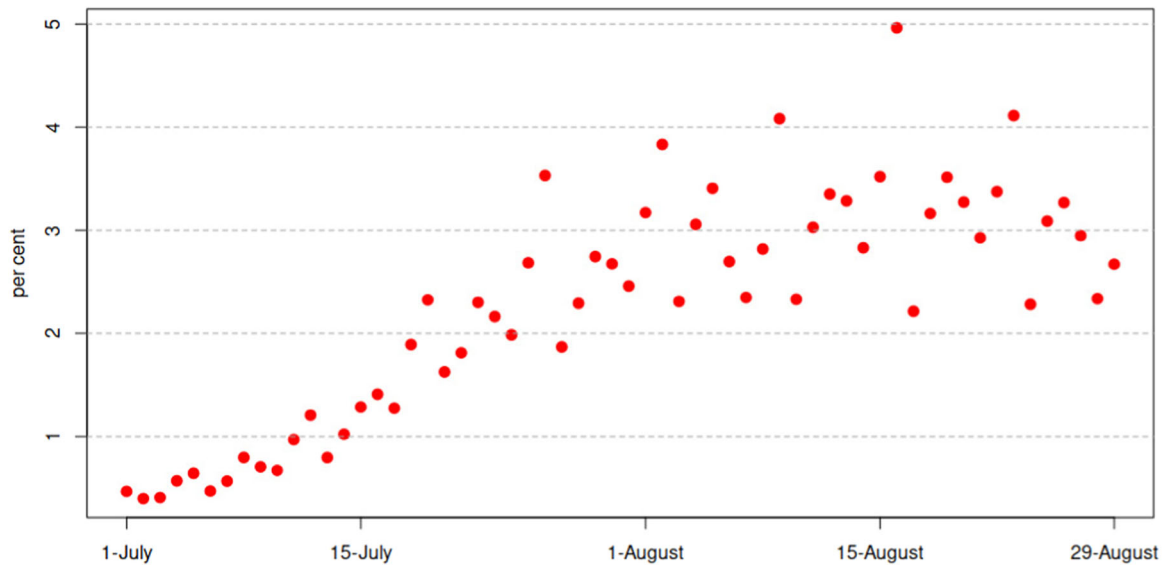

FIGURE 3 COVID-19 in Italy: Cases to swabs ratio (percent)

faith. We already observed several decisions based on black-boxes. We recall that restrictions were applied referring to the (effective) reproduction number R_t, whose misuse has been recently discussed.[5] More recently, a report (https://www.corriere.it/cronache/21_aprile_26/studio-segreto-riaperture-morti-stabili-fino-15-luglio-62a4ee8a-a6cd-11eb-b37e-07dee681b819.shtml) showed several scenarios where reopening daily activities, without restrictions, would lead up to 1300 deaths per day, with 200 as a lower bound, by mid of July. Well, on July 15, we recorded seven deaths. None of those scenarios was plausible and, thus, all were useless. Again, a further example, also cited by prime minister Mario Draghi in a

recent press conference, that sheds clouds on the way decisions are taken.

All these misleading information may arise only for one main reason: the models used to generate predictions, scenarios, and projections, are unreliable. They do not even fit well the data at hand. Sometimes we should remember that statistics can be a great worker but an evil mistress.

As discussed on the National Institute of Health webpage (https://www.iss.it/primo-piano/-/asset_publisher/3f4alMwzN1Z7/content/previsioni-scenari-proiezioni-come-si-anticipa-l-andamento-dell-epidemia), most analyses are based on a theoretical model without a full check that the model's assumptions are fulfilled or

"reasonable" at least. That may lead, sometimes, to biased inferences and results. We should prefer a data-driven approach. An example of how data-driven methods could be used is given by Pelagatti and Maranzano,[6] where the effects of restrictions during the second wave are assessed and different results from the official ones are estimated. Data versus modeling, the truth is in data. Mark Twain claimed, "People usually use statistics like a drunk streetlamp: more for support than for lighting," let us not fall into this error any further.

Fabio Divino[1]

Massimo Ciccozzi[2]

Alessio Farcomeni[3]

Giovanna Jona-Lasinio[4]

Gianfranco Lovison[5]

Antonello Maruotti[6,7]

[1]*Laboratory of Biostatistics and Computational Epidemiology, Department of Biosciences, University of Molise, Campobasso, Italy*
[2]*Medical Statistics and Epidemiology Unit, Università Campus Bio-Medico di Roma, Rome, Italy*
[3]*Department of Economics and Finance, University of Rome "Tor Vergata", Rome, Italy*
[4]*Department of Statistical Sciences, University of Rome "La Sapienza", Rome, Italy*
[5]*Department of Economics, Management, and Statistics, University of Palermo, Palermo, Italy*
[6]*Department GEPLI, Libera Università Maria Ss Assunta, Rome, Italy*
[7]*Department of Mathematics, University of Bergen, Bergen, Norway*

Correspondence

Massimo Ciccozzi, Medical Statistics and Epidemiology Unit, Università Campus Bio-Medico di Roma, Rome 00128, Italy.
Email: m.ciccozzi@unicampus.it

ORCID

Fabio Divino https://orcid.org/0000-0003-4107-3727
Massimo Ciccozzi http://orcid.org/0000-0003-3866-9239
Alessio Farcomeni https://orcid.org/0000-0002-7104-5826
Giovanna Jona-Lasinio https://orcid.org/0000-0001-8912-5018
Gianfranco Lovison https://orcid.org/0000-0003-3861-8204
Antonello Maruotti https://orcid.org/0000-0001-8377-9950

REFERENCES

1. Martinez EZ, Aragon DC, Nunes AA. Long-term forecasts of the COVID-19 epidemic: a dangerous idea. Revista da Sociedade Brasileira de Medicina Tropical. 2020;53. https://doi.org/10.1590/0037-8682-0481-2020
2. Agosto A, Campas A, Giudici P, Renda A. Monitoring COVID-19 contagion growth. *Stat Med.* 2021;40:4150-4160.
3. Alaimo Di Loro P, Divino F, Farcomeni A, et al. Nowcasting COVID-19 incidence indicators during the Italian first outbreak. *Stat Med.* 2021;40:3843-3864.
4. Girardi P, Greco L, Ventura L. Misspecified modeling of subsequent waves during COVID-19 outbreak: a change-point growth model. *Biometrical J.* 2021.
5. Maruotti A, Ciccozzi M, Divino F. On the misuse of the reproduction number in the COVID-19 surveillance system in Italy. *J Med Virol.* 2021;93:2569-2570.
6. Pelegatti M, Maranzano P. Assessing the effectiveness of the Italian risk-zones policy during the second wave of COVID-19. *Health Policy.* 2021;125:1188-1199.