Data inconsistencies do not allow valid modelling of the Italian COVID epidemic

Simone Cialdi [(1)], *Carlo La Vecchia* [(2)]

(1) Physics Department "Aldo Pontremoli", Università degli Studi di Milano, I-20133 Milano, Italia
(2) Department of Clinical Sciences and Community Health, Università degli Studi di Milano, I-20133 Milano, Italia

CORRESPONDING AUTHOR: Carlo La Vecchia, Deptartment of Clinical Sciences and Community Health, Università degli Studi di Milano, I-20133 Milano, Italia. E-mail: carlo.lavecchia@unimi.it

INTRODUCTION

We tried to monitor the epidemic of Covid-19 in Italy between February 22 and March 30 using the data disclosed by the Italian "Protezione Civile" (1).

Figure 1 gives the number of deaths, the cases in intensive care units (ICU) and the total cases of Covid-19 in Italy between February 22 and March 30. In order to show the changing velocity of the exponential diffusion of the disease, we used a logarithmic scale. When the disease is in the early stage its growth is exponential so the behavior in the log scale is linear and the velocity of the growth is represented by the slope of the line. When something happens during the diffusion of the epidemic the slope of the curve changes, for example due to the containment measure, the changing in the method to register the cases, the saturation of the places in the hospitals, and so on. In Figure 1 we can observe clearly the changes in the slope of the data and our goal is to find out the possible reasons for them. Moreover, in order to better indicate the slope we draw a line over the data.

Considering the curve of the total cases, there is a change in the slope of the curve between 26 February and 2 March. However, this is most likely due to the fact that in those days it was decided to make testing swabs only to symptomatic patients. This is a strong indication that the number of cases does not mirror the actual spread of the disease, but only the ability to monitor it.

Figure 2 considers cases in ICU and deaths in various areas of Italy. The number of cases admitted to ICUs in Lombardy shows a change in slope around 6 March. This appears to be due to the saturation of hospitalizations in Lombardy. Further, the curve of the deaths shows a clear change in slope around 9-10 March in Lombardy. Such a change of slope cannot be due to the containment measures taken since 8 March. Most likely, this change of slope (which refers to Lombardy only) reflects the fact that in the most affected areas (such as Bergamo) it was not possible to record all Covid 19 deaths. In this regard, the mayor of Bergamo stated there were 164 deaths in his town in the first two weeks of March this year, of which only 31 were attributed to the coronavirus. That compares with 56 deaths over the same period last year. Even adding the 31 coronavirus deaths to that total would leave *77* additional deaths, thus indicating that Covid 19 caused an appreciably larger number of deaths than officially recorded.

FIGURE 1. Number of deaths, the cases in the ICU (intensive care unit) and the total cases about the diffusion of the novel disease Covid-19 in Italy between February 22 and March 30.
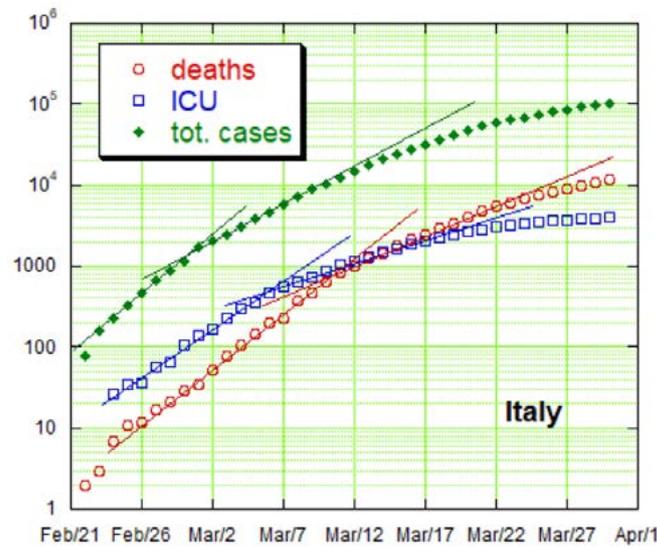
FIGURE 2. Comparison between cases in intensive care and deaths in four different parts of Italy.
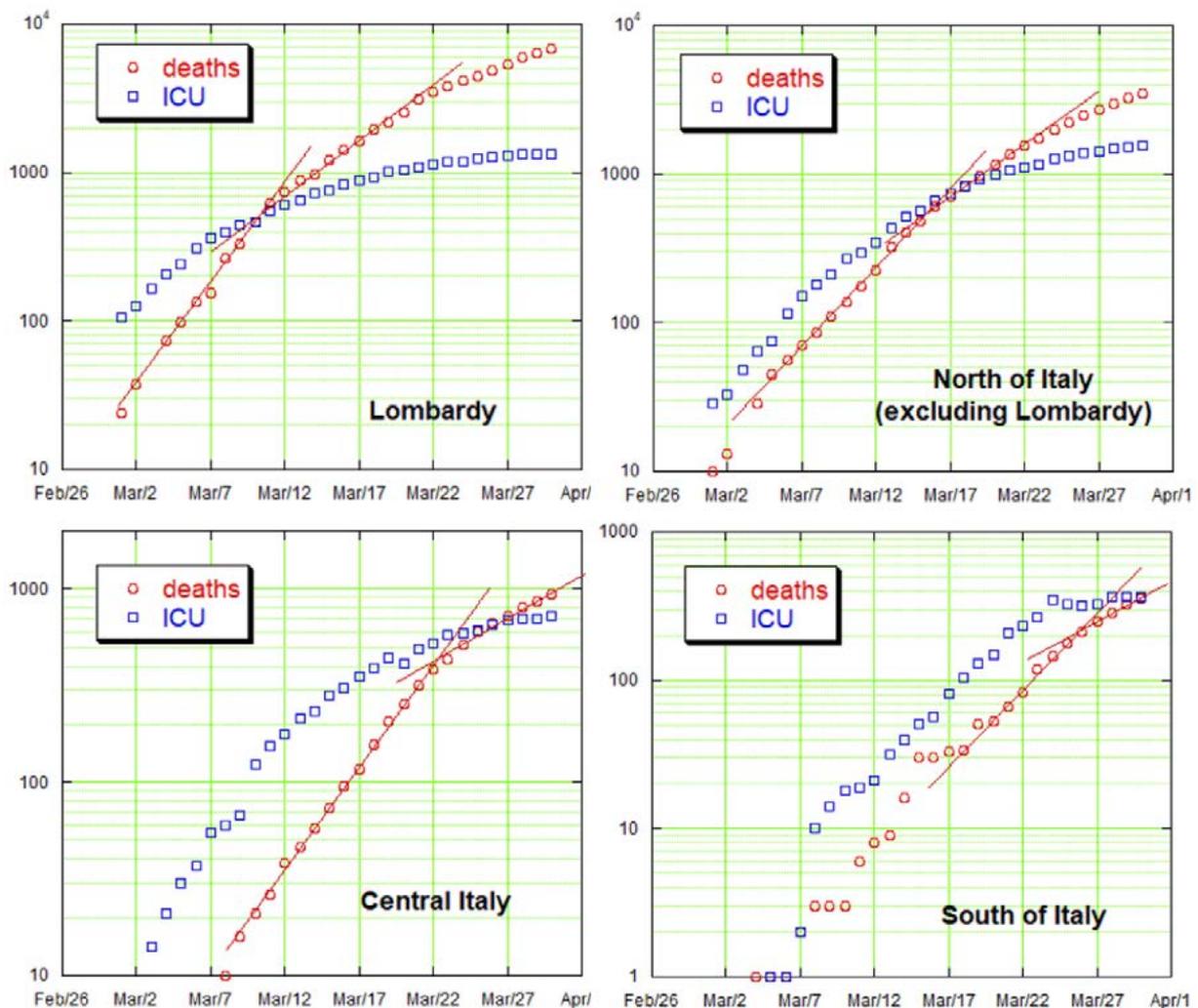
In order to better understand the changing slope of the deaths, we compared the cases in ICU and the deaths in four different parts of Italy (Figure 2): (1) Lombardy, (2) other northern Italy (excluding Lombardy, (3) central Italy, and (4) southern Italy. The death curves are cumulative, while the ICU curves refer to recorded cases in each day. There is a correlation between the point at which the death curves change in slope and the point at which the places in ICUs become comparable to the deaths.

This raises the doubt that the decrease in the slope of the deaths is at least partly due to the saturation of places in ICUs. In particular, in the early stage of the disease the rate of access to the ICU is very fast and the hospitalization time is long. Therefore, it is reasonable to observe this kind of correlation between the cases in the ICUs and the deaths. At later stages we can observe other changes in the slope of the deaths both in Lombardy and in the North of Italy. Part of this change could be due to containment measures, but considering the saturation of ICUs, the total number of deaths is probably underestimated, i.e., patients died at home without registration. It is, therefore, not clear how to use these data in a mathematical model to obtain a projection concerning the epidemic evolution.

In short, official data do not allow to consistently model, and hence project, the Italian Covid-19 epidemic. Before using these data in any statistical model, it is necessary to understand whether these slope changes are due to containment measures or other sources of error and bias. The changes in slopes of the curves is heavily correlated with the used methods to monitor the epidemic, and with the available places in ICUs. Thus, any prediction of a mathematical model will be distorted, and any prediction will be unreliable with available data.

References

1. https://github.com/pcm-dpc/COVID-19.

*