Title:

Comparing Length of Hospital Stay during COVID-19 Pandemic in the USA, Italy, and Germany

Running head:

Length Hospital Stay COVID-19

Authors:

Babak Jamshidi

*Corresponding author

Postdoctoral Researcher of Biostatistics and Epidemiology

Social Development and Health Promotion Research Center, Kermanshah University of Medical Sciences, Kermanshah, Iran

babak.j6668@gmail.com,

Shahriar Jamshidi Zargaran

Master of Science in Medical Engineering

Department of Medical Engineering, Tehran University of Medical Sciences, Tehran, Iran

shahriarjamshidy@gmail.com

Hakim Bekrizadeh

Assistant Professor of Statistics

Department of statistics, Payam-e-Noor University, Iran

bekrizadeh@pnu.ac.ir

Mansour Rezaei

Professor of Biostatistics

Social Development and Health Promotion Research Center, Kermanshah University of Medical Sciences, Kermanshah, Iran

mrezaei@kums.ac.ir

Farid Najafi

Professor of Epidemiology

Research Center for Environmental Determinants of Health, Kermanshah University of Medical Sciences, Kermanshah, Iran

farid_n32@yahoo.com

Title:

Comparing Length of Hospital Stay during COVID-19 Pandemic in the USA, Italy, and Germany

Running head:

Length Hospital Stay COVID-19

Abstract

Background

COVID-19 is the most informative pandemic in history. These unprecedented recorded data give rise to some novel concepts, discussions, and models. Macroscopic modeling of the period of hospitalization is one of these new issues.

Methods

Modeling of the lag between diagnosis and death is done by using two classes of macroscopic analytical methods: the correlation-based methods based on Pearson, Spearman, and Kendall correlation coefficients, and the logarithmic methods of two types. Also, we apply eight weighted average methods to smooth the time series before calculating the distance. We consider five lags with the least distance. All the computations are conducted on Matlab R2015b.

Results

The length of hospitalization for the fatal cases in the USA, Italy, and Germany are 2–10, 1–6, and 5–19 days, respectively. Overall, this length in the USA is two days more than in Italy and five days less than in Germany.

Conclusion

We take the distance between the diagnosis and death as the length of hospitalization. There is a negative association between the length of hospitalization and the case fatality rate. Therefore, the estimation of the length of hospitalization by using these macroscopic mathematical methods can be

introduced as an indicator to scale the success of the countries fighting the ongoing pandemic.

Keywords

Introduction

Up to the end of May 2020, over 6 million people have been infected by COVID-19, and around three-eighths of a million people have died of this infectious disease [1]. The epidemic COVID-19 is the most informative pandemic throughout history. These unprecedented recorded data give rise to some unprecedented concepts, discussions, and models. Modeling of the period of hospitalization by a macroscopic approach is one of these novel issues.

At first, throughout the paper, by the word "hospital", we mean:

"The hospital is an integral part of social and medical organization, the function of which is to provide for the population complete health care, both curative and preventive, and whose out-patient services reach out to the family in its home environment. The hospital is also a center for the training of health workers and bio-social research" [2].

Accordingly, all the hospitals serving whole COVID-19 services are acceptable regardless of the size, quality, and level of facilities.

Governments and health care authorities around the world are seeking evidence to evaluate their performances. They regularly explore the best implementation strategy for quality indicators and quantify the effect of quality indicators as a tool to improve the quality of hospital care [3]. Interest in comparative quality measurement and evaluation has grown considerably over the past three decades [4]. Introducing new scales improves both quality indicators and comparative quality measurements.

Undoubtedly, the length of hospitalization is a scale representing the function of the health system of countries. Normally, we associate more length of

hospitalization of fatal cases to the superiority of care for those admitted patients if the demographic and background variables of such patients are similar. The present paper is the first attempt to model the length of hospitalization for fatal cases of SARS-COV-2 using a macroscopic method.

There are two main approaches to study indicators like the length of hospitalization: microscopic and macroscopic. The microscopic approach is based on detailed data, while the macroscopic studies rely on a restricted piece of information –about the population, not about the individuals. Therefore, we classified the studies such as the systematic review of Rees et al. [5] and the 52 studies included in Table 1 [5] as microscopic. Our analysis is a macroscopic study because we base it on two available general variables of the intended countries: the number of confirmed cases and the number of deaths.

On the one side, the microscopic methods are more accurate and more reliable. If it is possible to conduct a detailed study like the cohort, we prefer the microscopic approach. On the other side, the macroscopic studies are faster and –both temporal and budgetary– more economical, therefore suitable for the countries poor in data. Particularly, this method is much more effective in case we want to conduct a correlational research project regarding some communities. For instance, to conduct the research entitled by the title of the present paper by a microscopic approach, we must collect data from the patients hospitalized in at least 20–30 hospitals in the three countries over a two or three-month period. It is noticeable that the key novelty of this study is adopting a macroscopic approach to deal with the characteristics of the pandemic.

For this study, we investigate the data from three major epicenters [6] with high-quality data collection systems: Germany, Italy, and the USA. In addition, they are remarkable from different points of view: the USA with the highest number of tests, confirmed cases, and deaths, Italy with the greatest case-fatality rate, and Germany with the most percentage of recovery worldwide [1]. Modeling the lag between the dates of diagnosis and death is done using two classes of macroscopic methods:

- The correlation-based methods based on Pearson [7], Spearman [8], and Kendall [9] correlation coefficients [10] and
- Three logarithmic methods of two types [11].

Applying the concept of cross-correlation to find the lag in periodic series has a long history. It is frequently applied in pattern recognition, single particle

analysis, electron tomography, averaging, cryptanalysis, and neurophysiology. The study of relationships between simultaneous time series, particularly those involving continuous human perceptions and performance, has been ongoing in many fields of medical sciences like psychology for several decades. Many researchers have applied the methods of the first class in medical fields to find the delay between the control and response variables [12–16].

There are different reports on the length of hospitalization in the countries of interest, and we use these studies to evaluate the accuracy of our correlational methods. Regarding Germany, the average time from the first symptoms to death was 14 days [17]. Also, the report of IHME (2020) showed that the length of hospitalization for fatal cases in Italy was 1–2 days less than in the USA, and the averages for both countries were around 10 days [18]. It has been estimated that the mean length of hospitalization for fatal cases in the USA was 15 days [19]. Besides, hospital stays lasted an average of 10.7 days for survivors and 13.7 days for non-survivors [20–21]. Finally, the median length of stay in the ICU has been reported to be approximately five days for patients who survive COVID-19, and six days for those who do not survive [22]. Considering three to four days for the distance between confirmation and transformation to the intensive care unit, the estimation of the length of hospitalization is between 9 and 10 days.

Finally, the questions that we address are:

- What are the findings from macroscopic analysis methods about the length of hospital stay in Italy, the USA, and Germany?
- Which macroscopic method is the most consistent with the reports?

Materials and Methods

The data are collected from the website Worldometer [1], and all calculations are done using Matlab R2015b.

Figure 1 illustrates the rationale behind the macroscopic method. The diagram shows that the cases that enter hospitals on a specific day will die $Lag - 2$ to $Lag + 2$ days later with the probabilities q_{n-2} to q_{n+2}, respectively.

<Figure 1>

Figure 1. The rationale behind the method

The cases that enter hospitals on the $(n - Lag)-$th day will die on one of the days $n - 2, n - 1,$ $n, n + 1,$ and $n + 2$ with the probabilities $q_{n-2}, q_{n-1}, q_n, q_{n+1},$ and q_{n+2}, respectively.

This paper is based on four presumptions:

> P.1. The changes in the number of new cases are reflected through the changes in the number of new deaths – or a modified form of it –after a lag (Figures 1, 2.A & 2.B).
> P.2. A fixed proportion of the cases go to the hospitals (Figure 2.C).
> P.3. A fixed proportion of the deaths occur inside the hospitals (Figure 2.D).
> P.4. The health care systems of Italy, the USA, and Germany are uniform.

<Figure 2>

Figure 2. Scatterplots to illustrate the relationship between the number of new cases and the smoothed number of new deaths in six days later in the USA, 4 March – May 31, 2020 (2.A), the relationship between the number of new cases and the number of new deaths in twelve days later in Germany, 13 March – May 31, 2020 (2.B), the relationship between the number of new cases and the number of new hospitalized cases in Minnesota, the USA, 6 March – December 23, 2020 (2.C), the relationship between the number of new deaths and the number of new deaths in the hospitals in England and Wales, Weeks 21 - 32, 2020 (2.D)

Initially, we need to justify the presumptions using some data as evidence. Unfortunately, It is not so simple to show the truth of Presumptions 2 and 3. Because most of the studies report the total and the current number of hospitalizations in which it is impossible to separate them into the proportions of discharges, recoveries, deaths, and new entrances. Moreover, almost all reports, aiming to find fatality per case, address the proportion of deaths in a group of cases admitted in some hospitals. Therefore, it is a worldwide lack of information on the published reports regarding the total count of deaths in the countries.

Scatterplots of Figure 2 justify the presumptions well: First, Subfigure 2.A represents that over the period 4 March – May 31, 2020, there is a linear regression

between the number of new cases and the smoothed numbers of new deaths on six days later in the USA [1]. The slope of the fitted line is around 0.07. Second, the scatterplot 2.B explains the linear relationship with slope 0.044 between the variables of new daily cases and new daily deaths in Germany from 13 March to 31 May 2020 [1]. Therefore,

$$\text{new cases} \propto \text{new deaths (after a lag)}.$$

Subfigure 2.C shows that from 6 March to 23 December 2020, there is a direct relationship between the number of new cases and the number of new hospitalized cases in Minnesota, the USA [23]. Almost 3% of the confirmed cases were admitted to hospitals. So,

$$\text{new cases} \propto \text{new hospitalizations.}$$

Finally, based on Subfigure 2.D, in England and Wales, between weeks 21 and 32 of 2020, around two-thirds of deaths involving COVID-19 occurred in hospitals [24]. Accordingly,

$$\text{new deaths} \propto \text{new deaths in hospitals.}$$

Combining all the above relationships, it is concluded that

$$\text{new hospitalizations} \propto \text{new deaths in hospitals (after a lag)},$$

which is the basis of our mathematical model.

Notice that since we can use smoothing methods, and these methods are able to collect and concentrate normally distributed data in the central points, it is possible to ignore the fourth presumption.

Consequently, modeling of the lag between diagnosis and death is done by using two classes of methods:

- The correlation-based methods based on Pearson [7], Spearman [8], and Kendall [9] correlation coefficients [10] and
- The logarithmic methods, including three methods [11].

Algorithm A presents the methods of the former class, and Algorithms B and C describe the first and second types of the latter class, respectively. The difference between Algorithms B and C is in the priority of standardization or division.

Whether Algorithm B defines the logarithmic method 1 or the logarithmic method 2 depends on the division by the standard deviation or mean of the deviations.

Algorithm A (the Pearson / Spearman / Kendall correlation-based method) [11]

```
A.1. i=1.

A.2. Consider the time series of the number of cases from
the first day and the number of deaths from the i-th day as
the variables.

A.3. Calculate the Pearson / Spearman / Kendall correlation
coefficient.

A.4. Calculate the square root of 1-|(A.3)| or one minus
the absolute value of the calculated correlation as the
distance scale.

A.5. i=i+1.

A.6. Return to A.2.

A.7. Select the i with the minimum value of distance as the
```

```
B.1. i=1.

B.2. Consider the time series of the number of cases from
the first day and the number of deaths from the i-th day as
the variables.

B.3. Standardize the time series of deaths and cases
through division by the standard deviation / mean of the
deviations.

B.4. Calculate the division of the standardized variables.

B.5. Calculate the mean of the logarithms of the ratios
(B.4) as the distance scale.

B.6. i=i+1.

B.7. Return to B.2.

B.8. Select the i with the minimum value of distance as the
desired lag.
```

Algorithm B (the logarithmic method 1 / 2) [11]

```
C.1. i=1.

C.2. Consider the time series of the number of cases from
the first day and the number of deaths from the i-th day as
the variables.

C.3. Calculate the division of the variables.

C.4. Standardize the ratio C.3 through division by the
standard deviation.

C.5. Calculate the mean of the logarithms of the
standardized ratios (C.4) as the distance scale.

C.6. i=i+1.

C.7. Return to C.2.

C.8. Select the i with the minimum value of distance as the
desired lag.
```

Algorithm C (the logarithmic method 3) [11]

To delete the noises and smooth the curves, we can use some weighted averages of
the 3, 5, or 7 nearest points based on:

- Uniform weights: $\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}$.

  (On Table 1, (3), (5), and (7) denote the average of 3, 5, and 7 nearest points, respectively)

- Geometric weights: $\frac{1}{1+2\sum_{i=2}^{n}\frac{1}{i}}(\frac{1}{n}, \frac{1}{n-1}, \ldots, \frac{1}{2}, 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n})$.

  (On Table 1, Geo. 3 and Geo. 5 denote the average calculated on the basis of geometric weights of 3 and 5 nearest points, respectively)

- Exponential weights: $\frac{1}{1+2\sum_{i=2}^{n}\frac{1}{i^n}}(1/i^n, \ldots, 1/i^2, \frac{1}{i}, 1, \frac{1}{i}, 1/i^2, \ldots, 1/i^n)$

  (On Table 1, (1/2).3, (1/2).5, (1/3).3, and (1/3).5 denote the average calculated from exponential weights founded on the base ½ of 3 and 5 nearest points, and the base 1/3 of 3 and 5 nearest points, respectively) [11].

It is noticeable that Algorithms B and C are applicable to positive series. Hence, we can apply them for the data of the USA, Italy, and Germany since dates 4 March, 22 February, and March 13, 2020, respectively. Similarly, we calculate the correlation-based scales for the pair of series after the aforementioned dates. After smoothing and applying the algorithms, we record the five fittest lags (Table 1).

Results

It is worth noting that "lag" means the number of days needed to be passed from diagnosis to death, and we study the lags 1 to 25. Table 1 summarizes the results of applying six methods for calculating similarity and nine smoothing methods (including real data).

<Table 1>

Table 1. The ranking based on the distance of the two series as a function for the lags 1–25

The correlation-based algorithm introduces 1–7, 1–5, and 5–14 days as candidates for the lag in the USA, Italy, and Germany, respectively (Table 1). The logarithmic methods 1 and 2 work similarly. They estimate the lags 4–8, 2–6, and 6–19 days as alternatives for the USA, Italy, and Germany, respectively. According to the third

logarithmic method, the lags 1–12, 1–12, and 6–19 days have the most probability to be the delay between two variables of the USA, Italy, and Germany, respectively. Overall, for the USA, Italy, and Germany, the most frequent lags in order of frequency are (6, 5, 7, 4, 8), (4, 5, 3, 6, 2), and (12, 11, 13, 10, 9), respectively. Generally, the calculated lags using the logarithmic methods are greater than the lags obtained by the correlation-based methods. Figures 3, 4, and 5 illustrate the comparison of the methods for scaling similarity and the regions under study.

<Figure 3>

Figure 3. The different distances between the number of cases and the number of deaths in the USA for lags 1–25 based on the nine smoothing methods by using the Pearson (3.A), Spearman (3.B), Kendall (3.C) correlation methods, and the first (3.D), second (3.E), and third (3.F) logarithmic methods

Figure 3 depicts the function of the different scales of similarity about the time series of the number of cases and the number of deaths in the USA. The comparison is made through the calculation of the distance of the lags 1 to 25. As it can be seen, the correlation-based methods work the same: an increasing trend with an eight-day cycle, and the maximum similarity between 5 and 6 days. The logarithmic methods 1 and 2 transform the desired lags to 6–7. Their plots are like the check mark. The plot of the correlation-based scales after a weighted averaging of 7 points is quite similar to the first two logarithmic methods. The third logarithmic method is N–shaped with the minimums at the beginning and around 10.

<Figure 4>

Figure 4. The distance between the number of cases and the number of deaths for lags 1–30 based on the nine smoothing methods for Italy (4.A), the USA (4.B), and Germany (4.C)

Figure 4 illustrates the different results of the third logarithmic method in the epicenters under study. In both of the plots of the USA and Italy, the lags 1 day, 9

days, and 10 days have the most probability. The lag plot of Germany is different from the other two countries. The mode is around 12, and the curve of the similarity is almost uniform, from a six-day lag to a thirty–day lag.

<Figure 5>

Figure 5. The frequency of the calculated lags between the confirmation and death for lags 1–25 days based on the nine smoothing methods and six methods for calculating similarity for Italy (5.A), the USA (5.B), and Germany (5.C)

Figure 5 shows that the USA and Italy have the same domain for the calculated lags. The difference between their bar plots is that the shape of Italy's one tends toward the left. Moreover, the patterns of the USA and Germany are similar, except for the six-day shift. It is observable that the histograms of Germany and the USA are approximately normal, but the graph of Italy is skewed.

Discussion

Statement of principal findings

Overall, for the USA, Italy, and Germany, the most frequent lags in order of frequency are (6, 5, 7, 4, 8), (4, 5, 3, 6, 2), and (12, 11, 13, 10, 9), respectively. The approximately normally distributed lags of the USA and Germany are similar except for the six-day shift, while the graph of Italy is skewed.

Strengths and limitations

Enormous factors are affecting the relationship between the number of confirmed cases and the number of deaths from a viral disease including the count of hospitalized cases, the quality of care in each country, the background risk of patients such as age and other health conditions, the preparedness of health systems, the ratio of patients to nursing staff, and the number of available ICU beds. We did not discuss these factors, and roughly assumed that the mentioned factors are overall the same for the countries under study. Our macroscopic analysis was based on two available general variables of the intended countries: the number of confirmed cases and the number of deaths. In comparison with detailed studies, the macroscopic studies are faster, –both temporal and budgetary– more economical, and less accurate. Therefore, this approach is suitable for countries

poor in data. Particularly, this method is much more effective in case we want to conduct a correlational research project regarding some communities.

Interpretation within the context of the wider literature

Considering the decreasing trend of the severity of the disease, conducting the reported studies before May 2020, the distance between the start of hospitalization and confirmation, and the statistical probabilities, the average of the lag between diagnosis and death for fatal cases in the USA, Italy, and Germany are 11–13, 9–12, and 15–18 days, respectively. In addition, to justify the skewness of the plot of Italy –in contrast to the normal plots of Germany and the USA-, the older age structure and the full-capacity and beyond-capacity attacks on healthcare systems in some regions may be helpful. Based on our model, the skewness in the plot of the number of deaths is also accompanied by skewness in the plot of cases; therefore, it does not affect our analyses.

Implications for policy, practice, and research

Since we can use smoothing methods, and these methods are able to collect and concentrate normally distributed data in the central points, it is possible to ignore the fourth presumption of the model.

The macroscopic studies are faster and –both temporal and budgetary- more economical therefore suitable for the countries poor in data. From a public health perspective, the new macroscopic method is insightful and helps the policy-makers compare their healthcare systems with those of the other involving countries and make some decisions including getting advice from the more successful countries. In this case, the Italian and American authorities may look for the reasons for the superiority –according to the introduced indicators- of the German system.

Finally, if there is a possibility to collect data from most patients and hospitals, and there is enough money and time available, it is preferable to adopt microscopic approaches to solve similar issues.

Conclusions

If we consider only the lag with the highest similarity, the function of the logarithmic methods 1 and 2 is better than the alternatives. Alternatively, if we take some lags into account or ignore the first day as the solution, the third logarithmic method works much better than the others. Finally, it is noticeable that we take the distance between diagnosis and death as the length of hospitalization. In addition, there is a negative association between the calculated length of hospitalization and case fatality rates. As of the end of May 2020, Germany, the USA, and Italy have case fatality rates $8605/183494 = 0.047$, $106432/1852029 = 0.057$, and $33415/232997 = 0.143$ percent, respectively. Therefore, the mathematical length of

hospitalization can be introduced as an index to scale the success of the countries fighting the ongoing pandemic.

Contributorship

BJ: Idea, Literature, Data, Methods, Programming, Interpretation, First draft. **HB:** Methods, Interpretation, Revision. **SJZ:** Data, Literature, Programming. **MR:** Conceptualization, Design, Final manuscript. **FN:** Conceptualization, Interpretation, Discussions, Revision.

Declaration of interest statement

We have no conflict of interest to declare.

Data availability

Availability of data	Sample statement
The number of confirmed cases and the number of deaths in	The data are available in [Worldometer], at https://www.worldometers.info/coronavirus/countries

Availability of data	Sample statement
Italy, the USA, and Germany	
Global epicenters of COVID-19	The data are available in [6], at https://doi.org/10.1515/em-2020-0009
The number of new cases and the number of new hospitalized cases in Minnesota, the USA	The data are available in [23], at https://www.health.state.mn.us/diseases/coronavirus/situation.html#raceeth1
The number of deaths involving COVID-19 occurred in the hospitals in the UK	The data are available in [24], at https://www.kingsfund.org.uk/publications/deaths-covid-19

References

1- Worldometer Website. https://www.worldometers.info/coronavirus/countries
2- World Health Organization, Technical report series No. 261. https://apps.who.int/iris/bitstream/handle/10665/40564/WHO_TRS_261.pdf?sequence=1&isAllowed=y
3- De Vos M, Graafmans W, Kooistra M et al. Using quality indicators to improve hospital care: a review of the literature. International Journal for Quality in Health Care 2009; 21(2): 119–129. https://doi.org/10.1093/intqhc/mzn059
4- Kelley ET, Arispe I, Holmes J. Beyond the initial indicators: lessons from the OECD Health Care Quality Indicators Project and the US National Healthcare Quality Report. International Journal for Quality in Health Care 2006; 18(1): 45–51. https://doi.org/10.1093/intqhc/mzl027
5- Rees EM, Nightingale ES, Jafari Y et al. COVID-19 length of hospital stay: a systematic review and data synthesis. BMC Medicine 2020;18: 270. https://doi.org/10.1186/s12916-020-01726-3

6- Jamshidi B, Rezaei M, Jamshidi Zargaran S et al. Mathematical modeling the epicenters of coronavirus disease-2019 (COVID-19) pandemic. Epidemiologic Methods 2020; 9*(s1),* 20200009. https://doi.org/10.1515/em-2020-0009

7- Pearson K. Notes on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London 1985; 58: 240–242.

8- Spearman C. The proof and measurement of association between two things. American Journal of Psychology 1904; 15 (1): 72-101. https://doi.org/10.2307/1412159

9- Kendall MG. Rank Correlation Methods. London: Griffin, 1970.

10- Jiaxing C, Kaow NY, Lu L et al. On triangular inequalities of correlation-based distances for gene expression profiles. bioRxiv, 2019.

11- Jamshidi B, Bekrizadeh H, Jamshidi Zargaran S et al. Some new methods to estimate the lag between two related variables. Preprint, 2020.

12- Gregson RAM. Time Series in Psychology. Hillsdale, New Jersey: Erlbaum, 1983.

13- Brunsdon TM, Skinner CJ. The analysis of dependencies between series in psychological experiments. British Journal of Mathematics and Statistical Psychology 1987; 40: 125–139.

14- Lindsey BG, Hernandez YM, Morris KF et al. Functional connectivity between brain stem midline neurons with respiratory-modulated firing rates. Journal of Neurophysiology 1992; 67: 890–904.

15- Pressing J. The referential dynamics of cognition and action. Psychological Review 1999; 106: 714–747. https://doi.org/10.1037/0033-295X

16- Holdefer RN, Miller LE, Chen LL et al. Functional connectivity between cerebellum and primary motor cortex in the awake monkey. Journal of Neurophysiology 2000; 84: 585–90.

17- Klabunde T, Giegerich C. How high and long will the COVID-19 wave be? A data-driven approach to model and predict the COVID-19 epidemic and the required capacity for the German health system. Preprint, 2020. https://www.researchgate.net/publication/340722893_How_high_and_long_will_the_COVID-19_wave_be_A_data-driven...

18- IHME COVID-19 health service utilization forecasting team. Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator days, and deaths by US state in the next 4 months. Preprint, 2020. http://www.healthdata.org/sites/default/files/files/research_articles/2020/covid_paper_MEDRXIV-2020-043752v1-Murray.pdf

19- Lewnard JA, Liu VX, Jackson ML et al. Incidence, clinical outcomes, and transmission dynamics of severe coronavirus disease 2019 in California and Washington: prospective cohort study. BMJ 2020; 369: m1923. http://dx.doi.org/10.1136/bmj.m1923

20- Coronavirus Resource Center. Johns Hopkins University & Medicine. Updated June 1, 2020. https://coronavirus.jhu.edu/map.html

21- Manke K. Long hospital stays, high rates of ICU admission for U.S. COVID-19 patients. Berkeley News 2020: https://news.berkeley.edu/2020/05/26/long-hospital-stays-high-rates-of-icu-admission-for-u-s-covid-19-patients/

22- ICNARC. ICNARC report on COVID-19 in critical care. London: Intensive Care National Audit and Research Centre, 17 April 2020.

23- Minnesota Department of Health Website. https://www.health.state.mn.us/diseases/coronavirus/situation.html#raceeth1
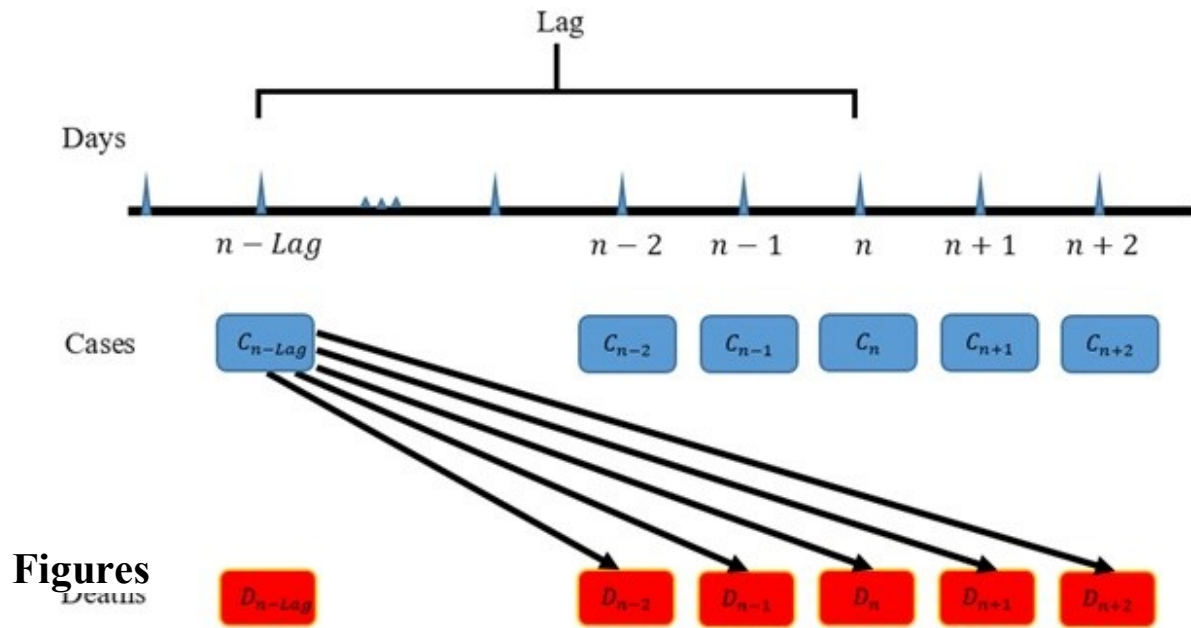
24- The King's Fund Website. https://www.kingsfund.org.uk/publications/deaths-covid-19

Tables

Table 1. The ranking based on the distance of the two series as a function for the lags 1-25

Smoothing	Country	USA					Italy					Germany				
	Ranking	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Smoothing	Method	Lag (Distance)														
(1)	Pearson	6	5	4	7	1	5	3	2	1	4	12	11	13	19	10
	Spearman	6	5	7	4	1	3	2	4	5	1	12	11	13	6	5
	Kendall	6	5	7	4	1	2	3	4	1	5	12	13	11	6	5
	Log1	6	5	7	4	8	4	3	5	6	2	12	13	7	8	11
	Log2	6	5	7	4	8	4	3	6	5	2	12	13	7	8	6
	Log3	1	2	3	9	10	3	2	1	9	10	12	13	8	7	19
(3)	Pearson	5	6	4	7	3	3	4	2	5	1	12	11	13	10	14
	Spearman	6	5	4	7	1	3	2	4	1	5	12	11	13	10	7
	Kendall	6	5	7	4	1	3	2	4	1	5	12	13	11	6	7
	Log1	6	5	7	4	8	4	5	3	6	2	12	13	11	10	9
	Log2	6	5	7	4	8	4	5	3	6	2	12	11	19	10	13
	Log3	1	2	9	10	11	1	2	10	9	11	12	11	10	6	18
(5)	Pearson	5	4	6	3	7	3	2	4	1	5	12	11	13	10	14
	Spearman	5	6	4	7	3	3	2	4	1	5	11	12	10	9	13

	Kendall	5	6	4	7	3	3	2	4	1	5	12	11	10	13	9
	Log1	6	5	7	4	8	4	5	3	2	6	10	11	9	12	19
	Log2	6	5	7	4	8	4	5	3	2	6	10	9	11	19	12
	Log3	1	2	10	9	11	1	2	10	11	9	11	16	15	17	6
(7)	Pearson	4	5	3	6	2	3	2	4	1	5	12	13	11	14	10
	Spearman	4	5	3	6	2	3	4	2	1	5	12	13	11	14	10
	Kendall	4	3	5	2	6	3	4	2	5	1	11	12	10	13	9
	Log1	6	5	7	4	3	4	3	5	2	6	10	17	16	9	11
	Log2	6	5	7	4	8	4	3	5	2	6	10	9	17	16	11
	Log3	1	2	10	11	12	1	10	11	9	12	15	14	6	16	12
Geo.3	Pearson	5	6	4	7	3	3	4	2	1	5	12	11	13	10	14
	Spearman	5	6	4	7	3	3	4	2	1	5	12	11	13	10	9
	Kendall	5	6	4	7	1	3	4	2	1	5	12	13	11	10	9
	Log1	6	5	7	4	8	4	5	3	6	2	10	11	12	9	13
	Log2	6	5	7	4	8	4	5	3	6	2	10	11	9	19	12
	Log3	1	2	10	9	11	1	2	10	9	11	11	12	13	17	18
(1/2).2 Geo. 2	Pearson	5	6	4	7	1	4	3	2	5	1	12	11	13	10	14
	Spearman	6	5	7	4	1	3	2	4	1	5	12	13	11	7	6
	Kendall	6	5	7	4	1	3	4	2	1	5	12	13	11	7	8
	Log1	6	5	7	4	8	4	5	3	6	2	12	11	13	19	9
	Log2	6	5	7	4	8	4	5	3	6	2	12	13	19	11	8
	Log3	1	2	9	10	11	1	2	9	10	11	12	11	13	10	6
(1/2).3	Pearson	5	6	4	7	3	3	4	2	1	5	12	11	13	10	14
	Spearman	5	6	4	7	3	3	2	4	1	5	12	11	13	10	9
	Kendall	5	6	4	7	1	3	2	4	1	5	12	13	11	10	7
	Log1	6	5	7	4	8	4	5	3	6	2	11	12	10	13	9
	Log2	6	5	7	4	8	4	5	3	6	2	10	11	12	9	19
	Log3	1	2	10	9	11	1	2	10	9	11	11	12	10	13	18
(1/3).2	Pearson	5	6	4	7	1	4	3	2	5	1	12	11	13	10	14
	Spearman	6	5	7	4	1	3	2	4	1	5	12	13	11	6	7
	Kendall	6	5	7	4	1	3	4	2	1	5	12	13	11	6	7
	Log1	6	5	7	4	8	4	5	3	6	2	12	13	11	19	8
	Log2	6	5	7	4	8	4	5	3	6	2	19	12	13	11	8
	Log3	1	2	9	10	11	1	2	9	10	11	12	11	13	10	6
(1/3).3	Pearson	5	6	4	7	3	3	4	2	5	1	12	11	13	10	14
	Spearman	6	5	7	4	1	3	2	4	1	5	12	11	13	10	7
	Kendall	6	5	4	7	1	3	2	4	1	5	12	13	11	6	10
	Log1	6	5	7	4	8	4	5	3	6	2	12	13	11	19	9
	Log2	6	5	7	4	8	4	5	3	6	2	19	12	13	11	10

	Log3	1	2	9	10	11	1	2	10	9	11	11	12	13	10	6
Total	**Total**	**6**	**5**	**7**	**4**	**8**	**4**	**5**	**3**	**6**	**2**	**12**	**11**	**13**	**10**	**9**

Figures

$$C_{n-Lag} \propto q_{n-2}D_{n-2} + q_{n-1}D_{n-1} + q_nD_n + q_{n+1}D_{n+1} + q_{n+2}D_{n+2}$$

Figure 1. The rationale behind the method

The cases that enter hospitals on the $(n - Lag)$ −th day will die on one of the days $n - 2, n - 1,$ $n, n + 1,$ and $n + 2$ with the probabilities $q_{n-2}, q_{n-1}, q_n, q_{n+1},$ and q_{n+2}, respectively.

Figure 2. Scatterplots to illustrate the relationship between the number of new cases and the smoothed number of new deaths in six days later in the USA, 4 March – May 31, 2020 (2.A), the relationship between the number of new cases and the number of new deaths in twelve days later in Germany, 13 March – May 31, 2020 (2.B), the relationship between the number of new cases and the number of new hospitalized cases in Minnesota, the USA, 6 March – December 23, 2020 (2.C), the relationship between the number of new deaths and the number of new deaths in the hospitals in England and Wales, Weeks 21 - 32, 2020 (2.D)
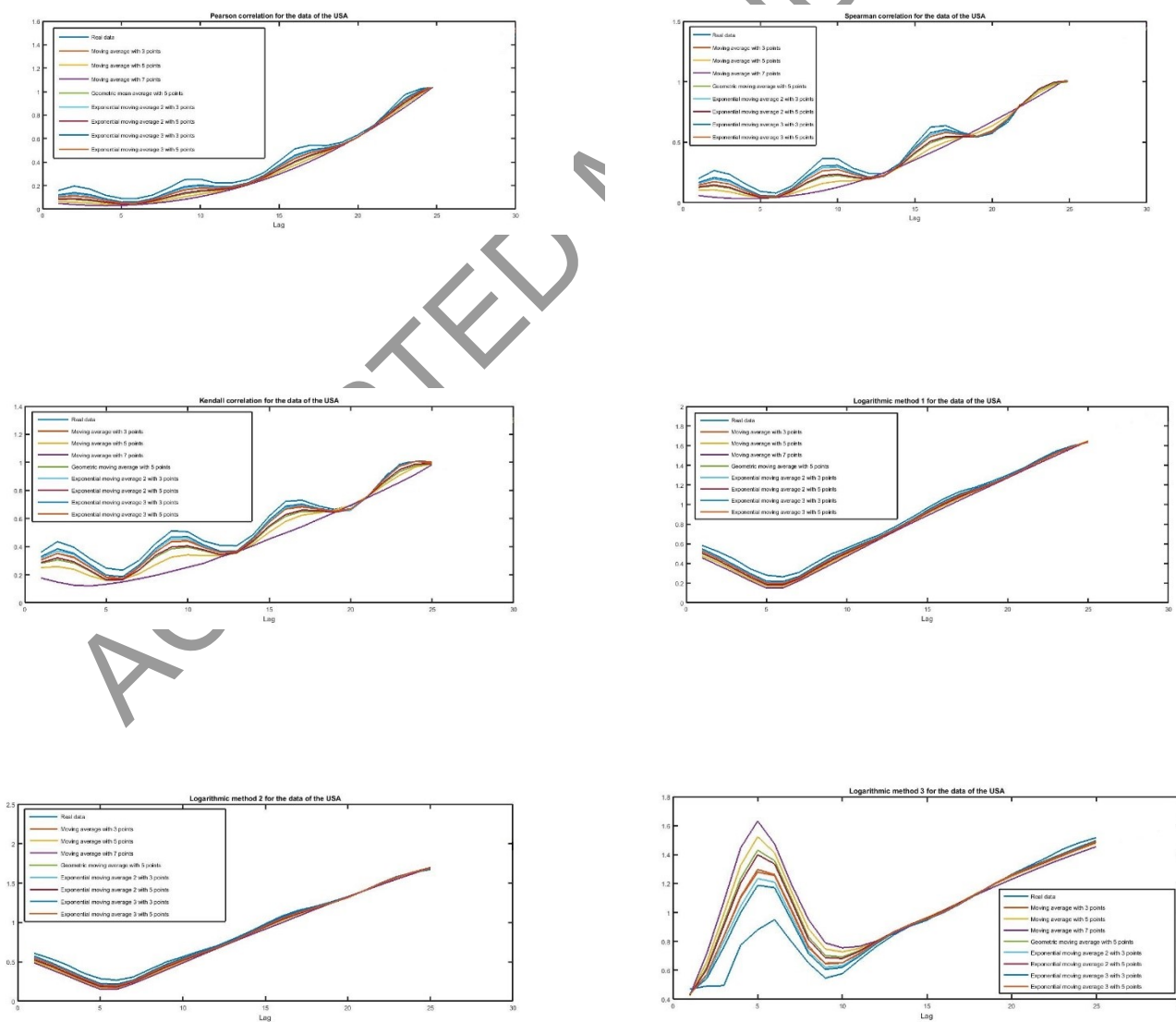
Figure 3. The different distances between the number of cases and the number of deaths in the USA for lags 1–25 based on the nine smoothing methods by using the Pearson (3.A), Spearman (3.B), Kendall (3.C) correlation methods, and the first (3.D), second (3.E), and third (3.F) logarithmic methods
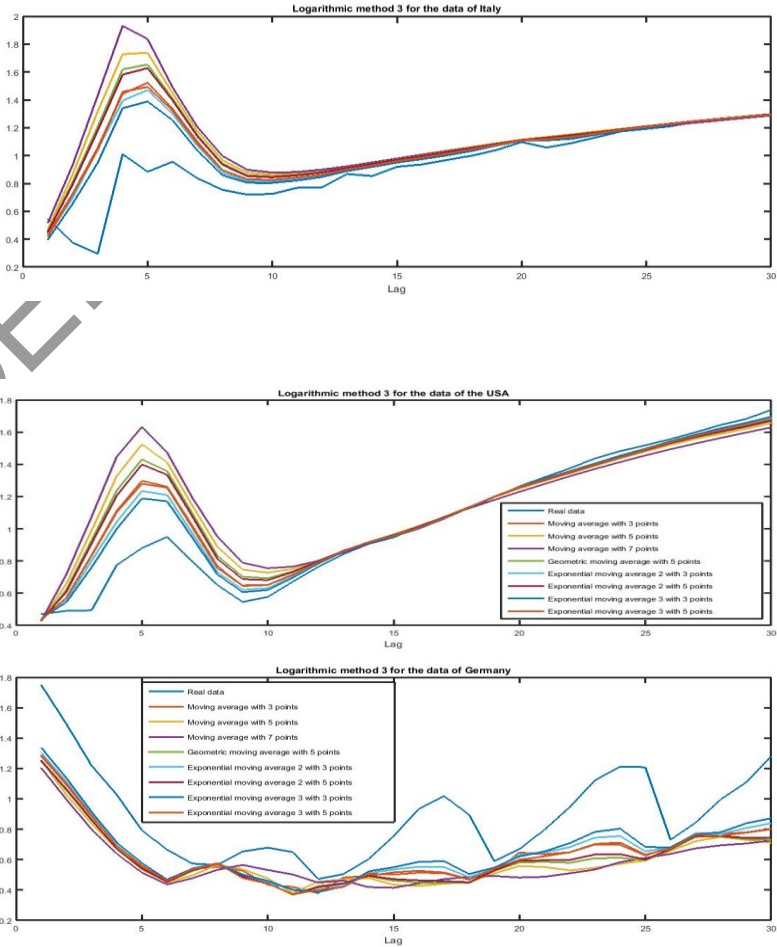
Figure 4. The distance between the number of cases and the number of deaths for lags 1–30 based on the nine smoothing methods for Italy (4.A), the USA (4.B), and Germany (4.C)
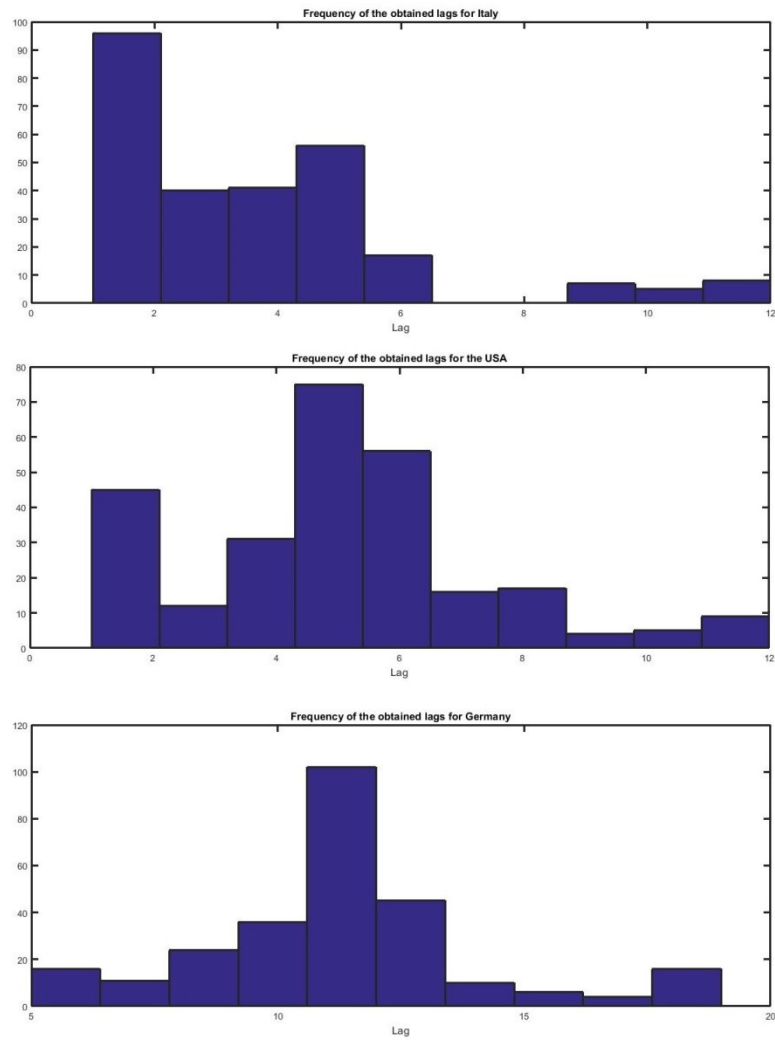
Figure 5. The frequency of the calculated lags between the confirmation and death for lags 1–25 days based on the nine smoothing methods and six methods for calculating similarity for Italy (5.A), the USA (5.B), and Germany (5.C)