Short paper

Joinpoint Regression to Determine the Impact of COVID-19 on Mortality in Europe: A Longitudinal Analysis From 2000 to 2020 in 27 Countries

Alessandro Rovetta

Correspondence

Alessandro Rovetta

R&C Research, Bovezzo (BS), Italy

ORCID: 0000-0002-4634-279X

Email: rovetta.mresearch@gmail.com

Telephone: +39 3927112808

Patamu Deposit Number: 172548

Abstract

The novel coronavirus disease 2019 (COVID-19) represented the most extensive health emergency in human history. However, to date, there is still a lot of uncertainty about the exact death toll the pandemic has claimed. In particular, the number of official deaths could be vastly underestimated. Despite this, many conspirationists speculate that COVID-19 is not a dangerous disease. Therefore, in this manuscript, we use joinpoint regression analysis to estimate the impact of COVID-19 in 27 European countries by comparing annual mortality trends from 2000 to 2020. Furthermore, we provide accessible evidence even for a non-expert audience. In conclusion, our results estimate that COVID-19 increased the overall mortality in Europe by 10% ($P < .001$). In 16 out of 27 countries (59.3%), the excess mortality ranged from 7.4% to 18.5% (all $P < .007$). Comparison of the observed mortality distribution with the null counterfactual showed that the mortality increase was highly significant across Europe, even considering only the nations with minor 2020 increases (highest $P = .014$, lowest mean = 2.4).

Introduction

Coronavirus disease 2019 (COVID-19) is a human-to-human transmissible infectious disease that reached the size of a pandemic in early March 2020, causing more than 5.5 million official deaths worldwide in less than two years [1]. The damage caused by this new coronavirus has been catastrophic, but the harms in the counterfactual scenario of the absence of non-pharmaceutical measures, therapies, and vaccines would have been even more dramatic [2,3]. Furthermore, the official number of COVID-19 deaths is plausibly an underestimate of the real figure, given the poor testing capabilities in the initial phase [4,5]. For this reason, scientists have begun comparing mortality statistics from previous years with current ones to highlight the actual epidemiological impact of COVID-19. Nevertheless, such a comparison is far from simple. Specifically, real time series can present critical issues such as trends, seasonalities, and level shifts. Therefore, direct comparisons between pandemic data and previous year averages can be improper or even misleading. Moreover, the further problem is that of trend estimation: indeed, determining the beginning of one trend and the end of another is a process that requires a high number of iterations to find the statistically most significant fit. In this regard, the Division of Cancer Control and Population Sciences of the National Institutes of Health has developed free software - called Joinpoint - to search for the best linear subtrends within a timeseries [6]. In this paper, Joinpoint was used to compare the annual mortality rates of 27 European countries before and after COVID-19 (from 2000 to 2020). The purpose of the manuscript is to provide epidemiologically relevant data and conclusive proof of the danger of COVID-19. Indeed, various conspiracy hypotheses have argued that a significant number of patients died from other causes, even if tested positive to the novel coronavirus 2019 [7]. Since risk perception is strongly influenced by how information and images are presented, a simple and intuitive figure has been developed to show the results to the lay public [8]. Finally, although more detailed surveys (e.g., stratified by age groups and periods) have been conducted, a more straightforward approach can provide clearer evidence and require fewer assumptions, reducing the likelihood of interpretative errors.

Methods

Summary of the procedure

Annual mortality data from 2000 to 2020 for all European countries were collected from the website "The World Bank" and downloaded in ".xls" format [9]. Mortality for 2020 has been obtained from the "Eurostat" website [10]. The time series were plotted for an initial check for normality and absence of marked outliers and heteroskedasticity. After that, the "Joinpoint" software - provided by the "Division of Cancer Control & Population Science" of the National Institutes of Health (NIH) - was adopted to break the time series into linear subtrends [6]. The last subtrend found was then analyzed and confirmed by a graph check and a linear regression analysis performed with the "XLSTAT" tool for Microsoft Excel v.2112 [11]. In particular, the tool automatically quantitatively checks the assumptions of the linear regression. Finally, the residual between the model's prediction for 2020 and the observed 2020 value was calculated for each time series. The Grubb test was applied to verify whether the 2020 residual was out of the distribution of previous residuals [12]. Welch t-test was exploited to compare two mortality distributions: the observed one and the counterfactual centered in 0. The shape of the two distributions was assumed to be the same.

Statistical details

Joinpoint regression. Most of the settings have been left at their defaults. The changed settings are specified below: Type of variable = Crude Rate (Death rate), Log transformation = No {y=xb}, Independent Variable = Year. In some cases, denoted with *, we have forced the model to introduce at least one joinpoint to fit the last values of the timeseries better (** was used for two joinpoints). All the graphs of these analyzes are reported as supplementary material to allow the reader an independent evaluation [13]. The acronym JPR-i indicates the i number of joinpoints. The absence of trend was indicated with NT.

Linear regression. Ordinary least square linear regression from the XLSTAT package was used to model the annual mortality trend from the last joinpoint through 2019. The standard assumptions of the model - i.e., normality of residuals, absence of outliers, and homoscedasticity - were automatically verified by Shapiro-Wilk and F-tests. The residual of 2020 was indicated with r_20 while the distribution of residuals from the last joinpoint up to 2019 with δ. The entire analysis is available as a supplementary Excel file to allow the reader an independent evaluation [13].

P-values. P-values were used as graded measures of the strength of evidence against the null hypothesis. Therefore we have not adopted dichotomic significance thresholds. However, we have divided the degrees of significance into low (P>.300), medium (.100<P<.300), fair (.050<P<.100), moderate (.010<P<.050), and high (P<.010). Indeed, we believe that this subdivision may be a compromise between the standard adoption of an a priori threshold - the purpose of which is to avoid interpretative biases - and the critical issues highlighted by Greenland et al. [14].

Results

The excess annual mortalities in 2020 of European nations (2020 %Exc) are shown in Table 1.

Nation	Outlier	2020 %Exc	SE (%Exc)	r_20/SD(δ)	Grubbs P	Trends	JP Years
Austria	2020	10.9	0.5	5.1	0.002	NT	NA
Belgium	2020	14.4	0.6	8.3	0.000	JPR1/NT*	2007
Bulgaria	2020	16.2	2.2	10.8	0.000	JPR-0	NA
Croatia	2020	7.0	6.9	2.9	>0.300	JPR-1*	2013
Cyprus	2017	0.4	1.3	1.1	>0.300	JPR-1	2015
Czech Republic	2020	16.4	0.6	7.7	0.000	NT	NA
Denmark	2018	-0.4	2.9	-0.4	0.059	JPR-1	2014
Estonia	2011	1.8	0.4	1.7	>0.300	JPR-2/NT	2007; 2011
EUROPE	2020	10.0	2.4	8.1	0.000	JPR-1	2008
Finland	2010	0.5	1.1	0.7	>0.300	JPR-1	2005
France	2020	8.8	1.9	8.6	0.000	JPR-1#	2007
Germany	2014	2.5	2.2	2.0	>0.300	JPR-1	2007
Greece	2012	2.6	4.4	1.2	0.000	JPR-1	2010
Hungary	2020	10.4	0.4	6.2	0.000	NT	NA
Ireland	2010	1.5	0.4	1.0	0.072	JPR-2/NT	2002; 2008
Italy	2020	18.5	3.9	9.2	0.000	JPR-1*	2006
Latvia	2019	2.2	2.4	1.8	>0.300	JPR-2	2006; 2009
Lithuania	2020	12.5	0.7	5.5	0.002	JPR-1/NT	2006
Luxembourg	2020	4.8	0.7	2.8	>0.300	JPR-1/NT*	2014
Malta	2012	3.5	0.7	1.2	>0.300	NT	NA
Netherlands	2020	7.8	2.3	6.9	0.000	JPR-1	2009
Poland	2020	13.4	4.8	9.3	0.000	JPR-1*	2014
Portugal	2020	7.4	6.5	6.0	0.007	JPR-1	2011
Romania	2020	13.2	2.8	9.2	0.000	JPR-2**	2007
Slovak Republic	2020	10.4	0.4	6.6	0.000	NT	NA
Slovenia	2020	13.5	7.4	10.2	0.000	JPR-1	2011
Spain	2020	13.0	5.1	5.9	0.004	JPR-1	2010
Sweden	2020	8.4	1.9	6.4	0.000	JPR-0	NA

Table 1. Excess annual mortalities in 2020 in all European countries. SE = standard error.

The general European situation is shown in Figure 1.

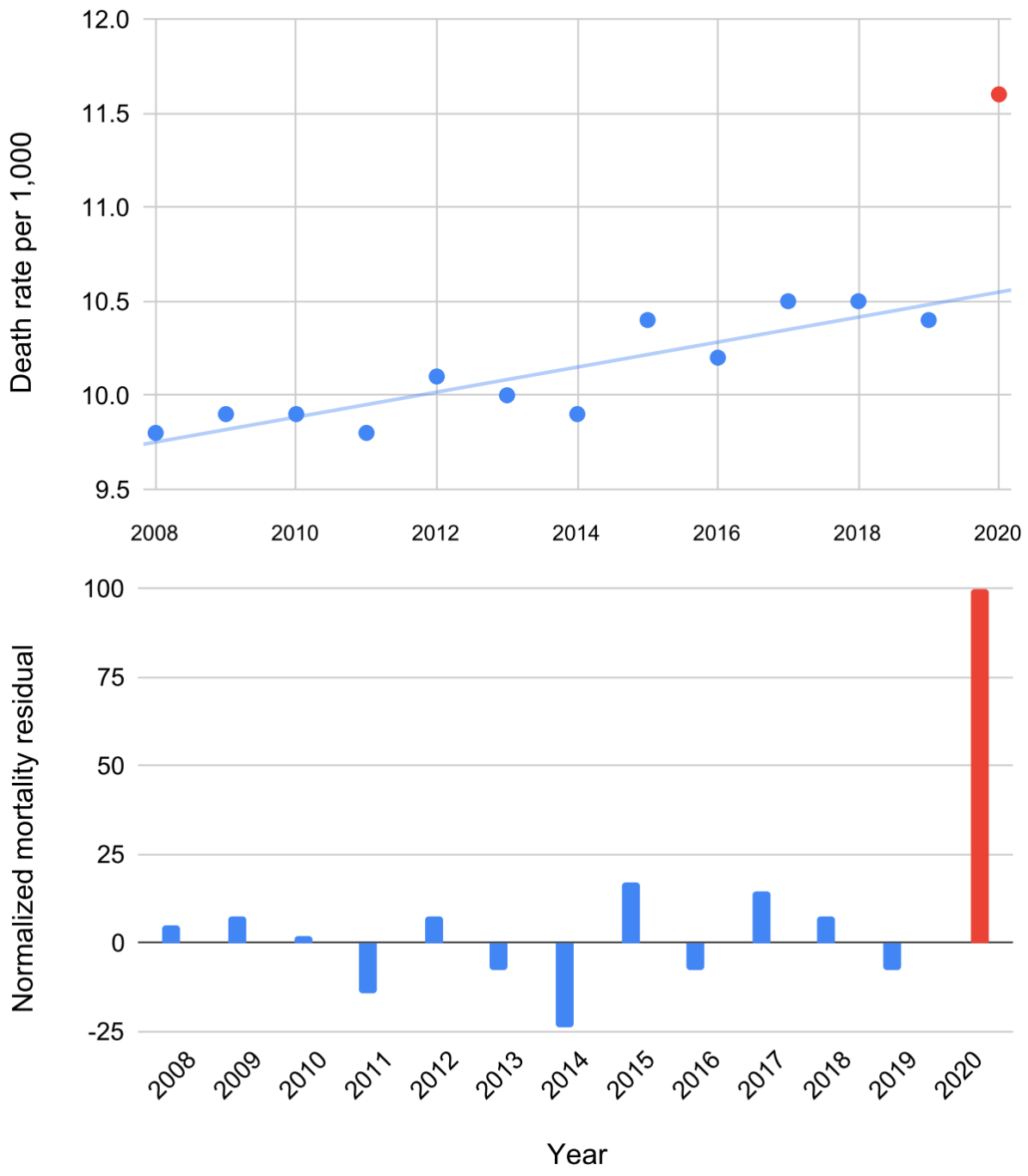

Figure 1. Differences between model predictions and observed death rates from 2008 to 2020 (residuals) in Europe. The abnormal increase in mortality during 2020 is visible to the naked eye.

The Grabb test identified anomalous excess mortality during 2020 in 66.7% of cases (18/27 EU countries). A high statistical significance ($P < .007$) was achieved in 59.3% of cases (16/27 EU countries). The percentage death rate excess ranged from 7.4% to 18.5% in the latter. However, all countries except Denmark recorded a positive death rate surplus during 2020 (mean = 8.2, SD = 5.7). Comparing this distribution with that translated into 0, we obtain a Welch $t = 5.3$ ($P<.001$). Repeating the calculation excluding the Grabb test results with $P < .05$, $t = 2.7$ ($P = .014$) has been found. Therefore, the increase in annual mortality was highly significant across the whole of Europe in 2020. Finally, the analysis revealed an increasing mortality trend in Europe over the last decade in 20 out of 27 countries (74.1%). The overall trend can be observed in Figure 1.

Discussion

The increase in mortality in Europe during 2020 was globally highly significant and marked. A heavy outlier was identified during 2020 compared to the model's predictions in most countries. Furthermore, the comparison of the null counterfactual scenario (i.e., no mortality increase) with the observed one revealed a highly significant and not negligible rise even for countries in which the Grabb test did not find an outlier. These findings provide strong evidence in favor of the absence of previous endogenous reasons capable of explaining the excess deaths in Europe. Since a vast literature discusses the virological and epidemiological reasons why it is a dangerous disease [15], we can conclude that COVID-19 is the main reason for such a scenario. Although mortality increases with age and the infection severity is linked to a wide range of local factors and comorbidities [7], it is wrong to conclude that COVID-19 is not harmful to young people. In particular, i) long-COVID (long-term effects of COVID-19) is a condition not uncommon in the younger age groups [16], and ii) the uncontrolled spread favors the appearance of variants of concern, capable of being much more dangerous even for the youngest (e.g., Delta and Omicron) [17]. Alongside this, joinpoint regression revealed growing mortality trends in almost all of Europe. This fact should alert health authorities beyond COVID-19. In conclusion, the author of this paper wishes these results to be helpful to the scientific community to estimate the actual number of deaths due to COVID-19 and to the population to achieve an adequate risk perception. Furthermore, it must be noted that the health damage would have been numerous times greater without the precautions taken by the governments. The same analysis will be carried out in a future study by stratifying the results by age group and season, thus providing more relevant epidemiological details. Finally, the reasons for the growing European mortality must be investigated.

Acknowledgments

I thank Dr. Akshaya Srikanth Bhagavathula for introducing me to the joinpoint regression method. I also thank Dr. Lucia Castaldo for her support in writing this manuscript.

References

[1] World Health Organization. WHO Coronavirus (COVID-19) Dashboard. URL: https://covid19.who.int/ (Accessed January 18, 2022).

[2] Mendez-Brito A, El Bcheraoui C, Pozo-Martin F. Systematic review of empirical studies comparing the effectiveness of non-pharmaceutical interventions against COVID-19. J Infect. 2021 Sep;83(3):281-293. doi: 10.1016/j.jinf.2021.06.018. Epub 2021 Jun 20. PMID: 34161818; PMCID: PMC8214911.

[3] Harder T, Külper-Schiek W, Reda S, Treskova-Schwarzbach M, Koch J, Vygen-Bonnet S, Wichmann O. Effectiveness of COVID-19 vaccines against SARS-CoV-2 infection with the Delta (B.1.617.2) variant: second interim results of a living systematic review and meta-analysis, 1 January to 25 August 2021. Euro Surveill. 2021 Oct;26(41):2100920. doi: 10.2807/1560-7917.ES.2021.26.41.2100920. PMID: 34651577; PMCID: PMC8518304.

[4] Modi C, Böhm V, Ferraro S, Stein G, Seljak U. Estimating COVID-19 mortality in Italy early in the COVID-19 pandemic. Nat Commun. 2021 May 12;12(1):2729. doi: 10.1038/s41467-021-22944-0. PMID: 33980836; PMCID: PMC8115692.

[5] World Health Organization. The true death toll of COVID-19 Estimating global excess mortality. URL: https://www.who.int/data/stories/the-true-death-toll-of-covid-19-estimating-global-excess-mortality (Accessed January 18, 2022).

[6] Joinpoint Trend Analysis Software v.4.9.0.0. Division of Cancer Control & Population Science, National Institutes of Health (NIH). URL: https://surveillance.cancer.gov/joinpoint/

[7] Rovetta A, Bhagavathula AS. Dying From COVID-19 or With COVID-19: A Definitive Answer Through a Retrospective Analysis of Mortality in Italy. medRxiv 2021.12.22.21268212; doi: https://doi.org/10.1101/2021.12.22.21268212.

[8] Ancker JS, Senathirajah Y, Kukafka R, Starren JB. Design features of graphs in health risk communication: a systematic review. J Am Med Inform Assoc. 2006 Nov-Dec;13(6):608-18. doi: 10.1197/jamia.M2115. Epub 2006 Aug 23. PMID: 16929039; PMCID: PMC1656964.

[9] The World Bank. Death rate, crude (per 1,000 people). URL: https://data.worldbank.org/indicator/SP.DYN.CDRT.IN (Accessed December 29, 2021).

[10] URL: Eurostat. Population change - Demographic balance and crude rates at national level. URL: https://ec.europa.eu/eurostat/databrowser/view/DEMO_GIND__custom_60961/settings_1/table (Accessed December 29, 2021).

[11] XLSTAT Free v.2021.1. XLSTAT by Addinsoft. URL: https://www.xlstat.com/en/

[12] Outlier tests for the normal distribution. Control Freak, Control charts for individual data points. URL: https://contchart.com/outliers.aspx (Accessed December 29, 2021).

[13] GitHub. SupplementaryMaterial/SupplementaryMaterial.zip. https://github.com/alex4lp/SupplementaryMaterial/blob/main/SupplementaryMaterial.zip (Accessed: January 20, 2022).

[14] Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, Altman DG. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016 Apr;31(4):337-50. doi: 10.1007/s10654-016-0149-3. Epub 2016 May 21. PMID: 27209009; PMCID: PMC4877414.

[15] Hu B, Guo H, Zhou P, Shi ZL. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol. 2021 Mar;19(3):141-154. doi: 10.1038/s41579-020-00459-7. Epub 2020 Oct 6. PMID: 33024307; PMCID: PMC7537588.

[16] Behnood SA, Shafran R, Bennett SD, Zhang AXD, O'Mahoney LL, Stephenson TJ, Ladhani SN, De Stavola BL, Viner RM, Swann OV. Persistent symptoms following SARS-CoV-2 infection amongst children and young people: A meta-analysis of controlled and uncontrolled studies. J Infect. 2021 Nov 20:S0163-4453(21)00555-7. doi: 10.1016/j.jinf.2021.11.011. Epub ahead of print. PMID: 34813820; PMCID: PMC8604800.

[17] Funk AL, Florin TA, Kuppermann N, Tancredi DJ, Xie J, Kim K, Neuman MI, Ambroggio L, Plint AC, Mintegi S, Klassen TP, Salvadori MI, Malley R, Payne DC, Simon NJ, Yock-Corrales A, Nebhrajani JR, Chaudhari PP, Breslin KA, Finkelstein Y, Campos C, Bergmann KR, Bhatt M, Ahmad FA, Gardiner MA, Avva UR, Shah NP, Sartori LF, Sabhaney VJ, Caperell K, Navanandan N, Borland ML, Morris CR, Gangoiti I, Pavlicich V, Kannikeswaran N, Lunoe MM, Rino PB, Kam AJ, Cherry JC, Rogers AJ, Chong SL, Palumbo L, Angelats CM, Morrison AK, Kwok MY, Becker SM, Dixon AC, Poonai N, Eckerle M, Wassem M, Dalziel SR, Freedman SB; Pediatric Emergency Research Network-COVID-19 Study Team. Outcomes of SARS-CoV-2-Positive Youths Tested in Emergency Departments: The Global PERN-COVID-19 Study. JAMA Netw Open. 2022 Jan 4;5(1):e2142322. doi: 10.1001/jamanetworkopen.2021.42322. PMID: 35015063.