Discrepancies between preliminary and final COVID-19 mortality data—the case of Serbia

Marko Galjak[1]* & Ivan Marinković[1]
[1] Institute of Social Sciences, Belgrade, Serbia
*Corresponding author: galjak@gmail.com

Accepted Manuscript

Galjak, M., & Marinković, I. (2023). Discrepancies between preliminary and final COVID-19 mortality data—The case of Serbia. *Annals of Epidemiology*, *84*, 41–47. https://doi.org/10.1016/j.annepidem.2023.05.006

ABSTRACT

Since the start of the COVID-19 pandemic, countries have scrambled to set up data collection and dissemination pipelines. These data end up in various online datasets. While it has been evident from the beginning that these preliminary data produced by some countries are not very reliable, the mortality data from Serbia seemed particularly problematic. These data are included in all major COVID-19 databases and utilized for research purposes worldwide. Discrepancies were identified between the preliminary mortality data reported through the emergency necessitated system and the final mortality data generated by the regular vital statistics pipeline. The number of deaths due to COVID-19 in Serbia, as reported preliminarily, does not align with the final death toll, which is more than three times higher. To evaluate the impact of the problematic data on research, we identified databases that include these data and reviewed the articles that utilized them. Our literature review identified at least 86 studies that were impacted by these problematic data. We strongly advise researchers to disregard the preliminary COVID-19 mortality data from Serbia. If all-cause mortality data is available, any preliminary data should be validated using excess mortality.

Keywords: Mortality, SARS-CoV-2, Serbia, Vital Statistics, Preliminary Data

List of abbreviations:

CFR - case fatality rate

SORS - The Statistical Office of the Republic of Serbia

WHO – The World Health Organization

WoS - Web of Science

GS - Google Scholar

INTRODUCTION

Background

Since the start of the COVID-19 pandemic, countries have scrambled to set up data collection and dissemination pipelines to quickly produce and publish the data regarding the pandemic. These data include the number of COVID-related deaths. Countries started reporting the daily

number of COVID-19 deaths. Online databases that aggregated these daily data emerged quickly.

With its aging population,[1] healthcare system which requires a lot of investment,[2] Serbia was initially considered vulnerable to the impact of the pandemic. However, from the start of the pandemic in March 2020, the daily reported mortality data kept pointing to a different conclusion. Serbia had few COVID-related deaths when compared to the number of COVID-19 cases. Serbia's case fatality rate (CFR) remained below 0.95% at the end of 2020, and as of September 2022, the CFR has decreased further to 0.72%.[3] Based on the available data, Serbia's COVID-19 outcomes have been comparatively better than those of its neighbouring countries. In Europe, Serbia was among the best-performing nations, with fewer COVID-19 deaths per confirmed COVID-19 case even compared to Norway. In contrast, countries that share similar demographics and socioeconomic profiles with Serbia, such as Bosnia and Herzegovina and North Macedonia, had case fatality rates (CFRs) that were more than three times higher.[4] However, after considering excess mortality data,[5–7] it became evident that Serbia's COVID-19 outcomes were more comparable to those of neighbouring countries rather than those of Norway or Denmark.[8]

This study compares the preliminary COVID-19 mortality data from Serbia and the official vital statistics in order to determine the presence or absence of any discrepancies. We conducted a literature review of studies that use preliminary data on COVID-19 mortality in Serbia to determine the extent to which potentially suspect data is being employed in research. We discuss the implications of these findings, and provide recommendations for researchers encountering anomalous COVID-19 mortality data from Serbia in public datasets. We then provide general recommendations for handling potentially anomalous COVID-19 mortality data from other countries. This study aims to provide an analysis of the potential discrepancies between final official mortality data and preliminary daily reported data on COVID-19 mortality in Serbia, with the goal of contributing to the scientific community's understanding of the situation. We highlight the presence of problematic data that continues to be published in major databases and used widely in academic publications. In addition, we conduct a brief comparison between the preliminary and final data from other countries to determine the extent to which Serbia stands out and whether such discrepancies exist in other countries.

Material and methods

COVID-19 mortality data

We conducted a comparison between the aggregated preliminary, daily reported COVID-19 mortality data from Serbia and the final vital statistics data, which is considered more reliable, with the aim of evaluating potential differences between these two datasets. The comparison was possible only for the first two years of the pandemic – 2020 and 2021 - due to the vital statistics publishing schedule. We conducted a month-by-month comparison of the two datasets from March 2020 to the end of 2021. The comparison between the two different pipelines for gathering mortality data is illustrated in Figure 1.

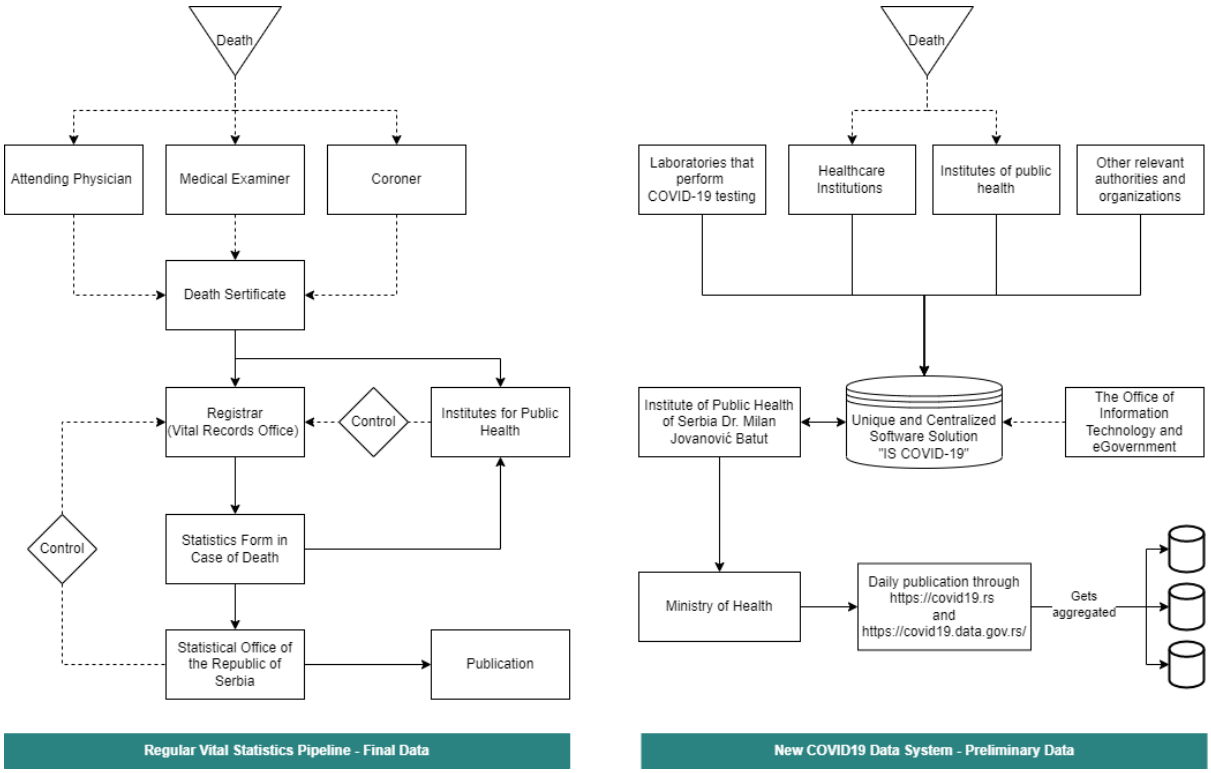

Figure 1: Final data – produced using a regular vital statistics pipeline (left), and the preliminary data – based on the new, emergency-necessitated system for producing COVID-19 data (right)

Daily Reported (Preliminarily) Data

The need for a faster distribution of pandemic-related data led to the development of a new centralized COVID-19 surveillance system, using a Unique and Centralized Software solution named 'IS COVID-19'. This Serbian Ministry of Health's database is run by the Institute of Public Health of Serbia – Dr Milan Jovanovic Batut and The Office of Information Technology and eGovernment.[9] The database is being used to report daily COVID-19 death count (among other COVID-19 statistics). These statistics are reported by government officials, distributed by the media, and aggregated into online COVID-19 databases. The first (daily reported) COVID-19 death in Serbia occurred on 20th March 2020.

The Vital Statistics (Final) Data

Gathering of vital statistics in Serbia is regulated by the Law on Official Statistics and Official Statistics Programme for 2021-2025 (and its earlier edition for 2016-2020).[10] As the exclusive provider of official government statistics, SORS is responsible for designing survey instruments and managing their implementation. Following rigorous quality control measures, including analysis and tabulation, the statistical data is released according to an established calendar and through designated publication channels.[10] The official mortality data for the preceding year is routinely released in July. As of the time of writing, the authors had access to the final mortality data for 2020 and 2021.[11] The detailed mortality dataset is available in the offline mortality database maintained by the World Health Organization (WHO), which currently includes data up to 2020.[12] The authors were able to obtain detailed mortality data for both 2020 and 2021 through a special request made to the SORS.[11]

Search strategy

To assess the potential impact of the preliminary data from Serbia on research, the authors conducted a review of published research papers that had utilized this data. We used a two-pronged approach. The first approach involved a direct examination of the sources of the data, specifically the Serbian government portals covid19.rs and covid19.data.gov.rs. The second approach examined databases that aggregate the preliminary data from Serbia.

The search period for the present study extended from March 2020, when preliminary data for Serbia began to be reported, to September 2022, at the time of writing this article.

The present study identified three main databases that contained preliminary data on COVID-19 mortality for Serbia: the Johns Hopkins database,[13] Our world in data,[4] and Worldometer.[14] While there are other databases that aggregate daily reports from Serbia, we chose these three for their span, accessibility, widespread use.

To identify research that may have been affected by the questionable preliminary mortality data, the authors conducted a search using two databases: Web of Science (WoS) and Google Scholar (GS). Appendix A contains a detailed description of the search strategy, including search strings used in the analysis.

The authors conducted the GS search using the Publish or Perish software[15] and used Zotero[16] to perform DOI lookup and export bibliographic data.

Eligibility

The study included research studies that explicitly focused on Serbia and used the preliminary data. In addition, cross-country studies that incorporated the preliminary COVID-19 mortality data from Serbia in their analyses were also included.

A significant number of research studies that cited one of the databases were easily excluded, such as those related to specific countries or regions that were not relevant to Serbia. Numerous research papers related to the pandemic cited the data from the databases in their introductory sections to provide overall figures concerning the pandemic, including the number of deaths. Research studies that reported on worldwide deaths caused by COVID-19 but did not include Serbia in their analyses were excluded. Although the worldwide COVID-19 deaths number includes the underreported numbers from Serbia, which technically could make the reported number incorrect, we assessed that this factor alone did not have a significant impact on the research. Due to the population size of Serbia, the potential underreporting could not have resulted in a significant difference in worldwide numbers. Nonetheless, we included research studies that reported the number of COVID-19 deaths specifically in Serbia due to the substantial absolute difference in reported numbers.

All studies that employed the preliminary COVID-19 mortality data from Serbia in statistical analysis were included, regardless of their methodology (even though some analyses used methodologies that were more resistant to outliers than others).

To determine if Serbia was included in the analysis, we checked whether it was listed in the main text, tables, figures, and supplementary data (where available). Several research studies utilized data from the affected databases but did not explicitly mention the countries that were included in their analyses. We excluded those studies unless they analysed a considerable number of countries, in which case it was reasonable to assume that Serbia was included in the analysis. The complete list of studies is provided in the Appendix B Table B.1 (research that

includes mortality data from Serbia) and Table B.2 (research that likely includes mortality data from Serbia).

Citations

We collected the number of times that each identified research study and paper was cited in both the WoS Core Collection and GS. The citation count was used as a criterion for selecting studies to address in the discussion section of this paper. The combined number of citations was calculated by summing the citation counts, without taking into account the number of distinct papers that referenced them.

RESULTS

Data Comparison

The final detailed mortality data, generated through the standard vital statistics process, provided a more comprehensive view and revealed significant discrepancies between the two datasets. According to vital statistics data, Serbia lost 38,098 people to COVID-19 in the years 2020 and 2021. However, the daily reported preliminary data indicated a much lower figure of only 12,717, which is nearly three times less than the final count.

Although it is common for final data to differ from preliminary data, the significant disparities between the two in this case are noteworthy. By examining monthly death tolls, it becomes apparent that the differences have fluctuated significantly over time, ranging from 170% to 520%. The considerable variability in the relative difference between the preliminary and final data prompts examination of the usefulness of the preliminary data for accurately assessing COVID-19 mortality at any given point in time.

The timing of these variations is noteworthy, particularly given that Serbia experienced an early easing of restrictions in May 2020, followed by national parliamentary elections in May and June that featured large, in-person rallies and intensive campaigning (Figure 2). Interestingly, June 2020 had the greatest disparity between the reported number of deaths and the actual number of deaths according to the vital statistics data.

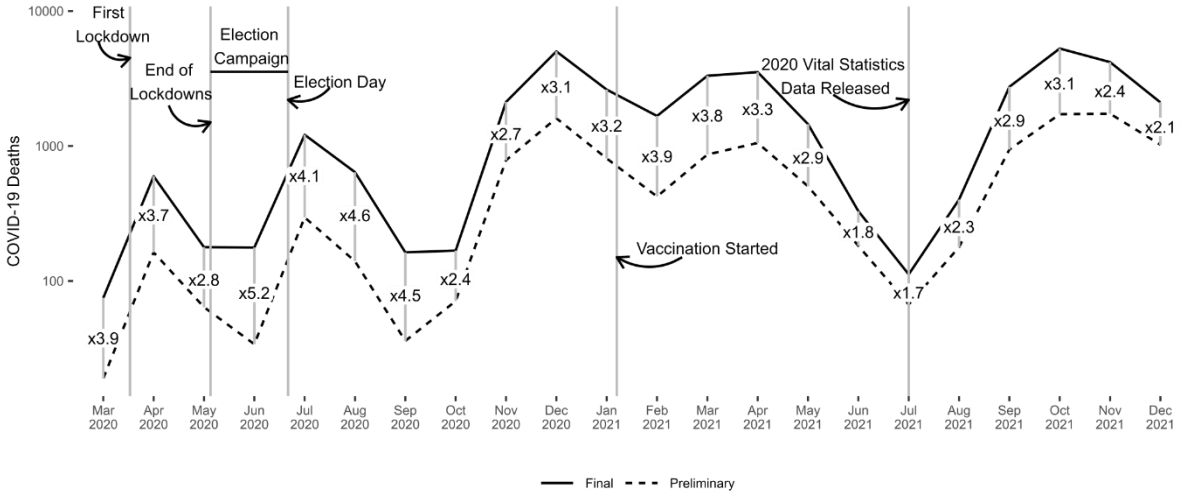

Figure 2: Monthly COVID-19 deaths in Serbia for 2020 and 2021 – preliminary and final data

x – multiple preliminary deaths (i.e., ratio showing how many times more people died from COVID-19 than preliminarily reported).

While this study did not extensively analyse the differences between preliminary and final data for other countries, we briefly compared these differences to those observed in Serbia. While variations were noted in other countries, they were not as pronounced as those seen in Serbia (Appendix D, Table D.1). Across 18 European countries, the final vital statistics data showed a mean increase of approximately 20% in COVID-19 deaths compared to the preliminary data.

Impact of Preliminary Data

Among the 675 studies that we reviewed on this topic, 96 studies were identified as potentially utilizing preliminary data from Serbia (Figure 3). Of these, 60 studies were published across 50 different journals indexed in the WoS, with 42 having an Impact Factor (μ=6.25, min=1.169, max=39.19). These included papers from 37 different WoS Categories. The category with the most papers was Public, Environmental & Occupational Health (n=15), followed by Environmental Sciences (n=12) and Multidisciplinary Sciences (n=8). Most of the research papers had little impact on citation (Me =3/μ=10.32). While the majority of the research papers had limited impact measured by the number of citations (Me =3/μ=10.32), some had significant influence, with two articles accruing over 100 citations. The identified research was cited 603 times in publications indexed by the WoS.

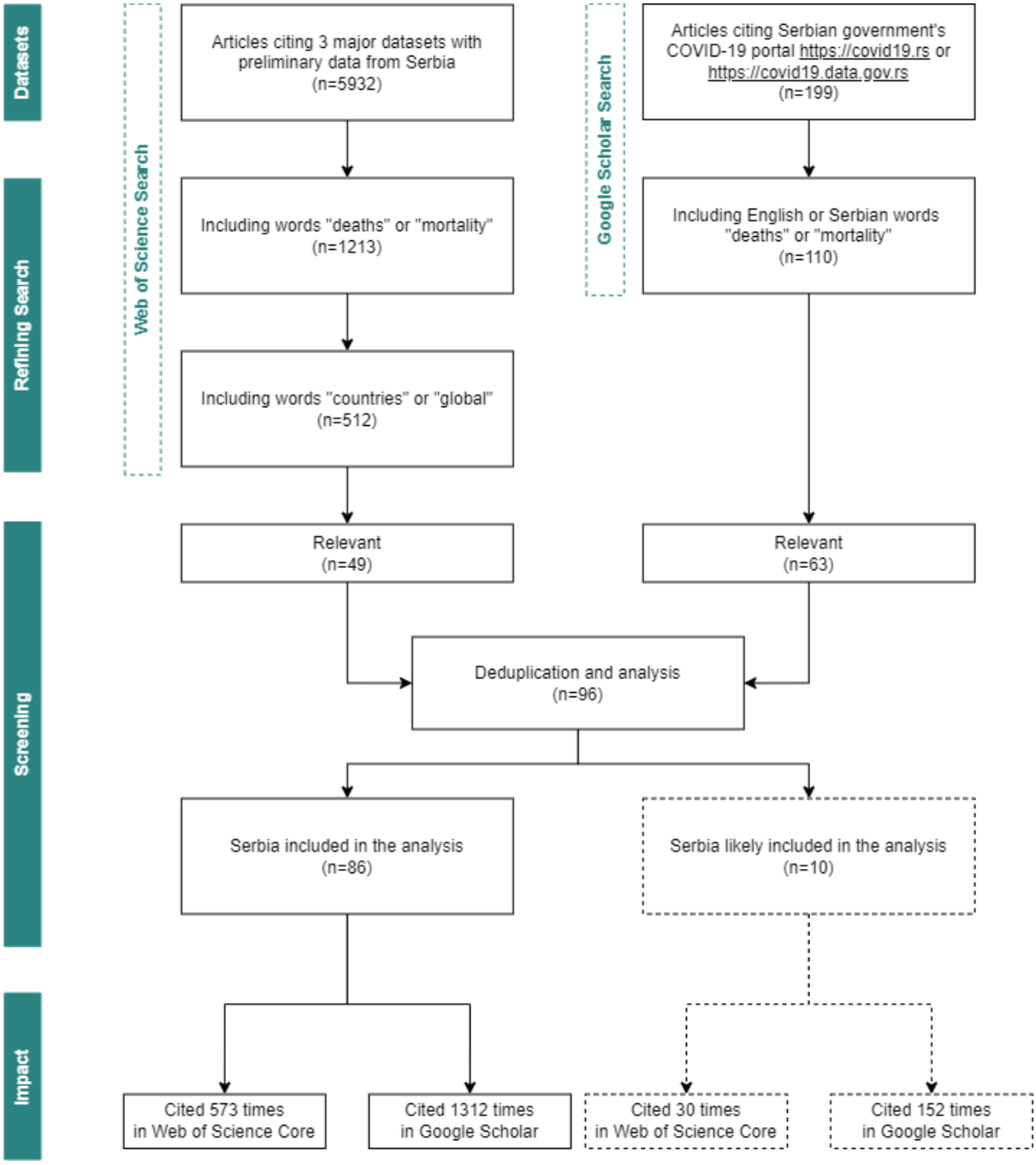

Figure 3: Process of finding the Serbian preliminary COVID-19 data being used in published research.

DISCUSSION

The vital statistics systems are considered the most reliable and definitive source of information on COVID-19 mortality.[17] The disparity between the preliminary and vital statistics data in Serbia, that we have demonstrated, is not what we commonly see in other countries. In the United States, the emergency system for tracking COVID-19 data named COVID Data Tracker, set up by The Centers for Disease Control and Prevention,[18] produces data that closely resemble the data coming from the National Vital Statistics System.[19]

The comparison between the preliminary and final data from Serbia suggests a significant underreporting of COVID-19 deaths in the preliminary data, with the preliminary numbers indicating about three times fewer COVID-related deaths than the final data. While this underscores the potential limitations of relying on preliminary data, it also emphasizes the importance of using validated, final data for research and decision-making.

The research we identified using the problematic preliminary data included a study investigating whether the BCG vaccine offered protection from COVID-19, which was one of the most cited papers.[20] In this study, the inclusion of Serbia in their analysis had a significant impact on the coefficient of determination. Such outliers could potentially lead to overinflated p-values or premature dismissal of a hypothesis without publication of the results.

The importance of high-quality data has not diminished as the pandemic has progressed. The public's confidence in government institutions proved to be a cornerstone of managing the COVID-19 crisis.[21,22] Any irregularities, or even the perception of irregularities, in data can erode public trust in government institutions, especially in the midst of a pandemic where public-health decisions about mitigation efforts need to be evidence-based. It is crucial to ensure the accuracy and transparency of data collection and reporting in order to maintain public trust and support for necessary public health measures. Underreporting can also contribute to a distorted perception of the severity of the pandemic among the public, potentially leading to underestimation of the risks and inadequate measures being taken to control the spread of the virus. In order to offer effective guidance on public health measures, Serbia requires a reliable epidemic intelligence system that can provide subnational epidemiological data for modelling disease transmission.[8]

Causes of the discrepancy

Can the differences in the definition of a case, that explain discrepancies between some countries,[23] also explain the discrepancies between the two datasets in Serbia? Our analysis suggests that the discrepancies between the two datasets in Serbia cannot be explained by differences in the definition of a case. While the technical details of the new system for collecting COVID-19 data are opaque, the available documents suggest the use of the same definition of a case as outlined by the WHO. The mortality dataset from the vital statistics provides information on two causes of death related to COVID-19: where the virus was confirmed (U07.1) and where its presence was suspected (U07.2). Out of all COVID-19 deaths, 92.7% (35,287) were confirmed cases, while only 7.3% (2,771) were suspected cases.

Others have highlighted the incongruities between the COVID-19 preliminary data and the excess deaths reported by the SORS,[24,25] and this topic of undercounting has also been covered by the media.[26] However, the causes of this discrepancy at the time of writing this article are yet unclear and require further research.

What about other countries?

While we need to recognize the importance of the newly established mortality data pipelines, it is also important to note that the vital statistics systems are the best source of information on COVID-19 mortality.[17] The differences between the preliminary and vital statistics data in Serbia appear to be more pronounced than in other countries. In the United States, the emergency system for tracking COVID-19 data, set up by The Centers for Disease Control and

Prevention, produces data that largely corresponds to the data coming from National Vital Statistics System.[19] The UK Office for National Statistics' analysis of COVID-19 mortality data has been conducted more rapidly than in many other countries, where the data release processes can be lengthier.[27]

During 2020, some countries faced concerns about the accuracy and reliability of their daily mortality data related to COVID-19. In certain countries, researchers raised questions about the quality of data on COVID-19 deaths, and there were even allegations of possible manipulation or poor data quality.[28,29] Brazil[29–31] and India[32] were among the countries where these concerns were most notable. One study indicated that authoritarian regimes were likelier to manipulate COVID-19 data.[33]

Detailed vital statistics data has been released in some countries, and in these cases, differences between the preliminary COVID-19 data and the vital statistics have been observed. Nevertheless, the disparities observed in Serbia are far greater than those reported in other countries, suggesting that there may be additional factors contributing to the discrepancy. There are two major consequences of the problematic preliminary data. Firstly, it raises concerns about the reliability of any preliminary mortality data and the degree of trust that can be placed in it in the time of crisis. Secondly, it can significantly change the outcomes of studies relying on it, which highlights the need for a standardized approach to handle outliers in publicly available COVID-19 mortality datasets.

Recommendations

As a suggestion to database curators, we recommend the removal of Serbian daily reported mortality data, as well as any other outliers that may significantly affect the reliability of the dataset. For example, French Institut national d'études démographiques database[34] does not include the Serbian data.

In big datasets that encompass nearly all countries, outliers are inevitable. The preliminary COVID-19 mortality count from Serbia is more than just an outlier; it represents an extreme value that will skew the results of many analyses. We argue that researchers should be aware of these outliers and deal with them accordingly, either by eliminating them, or finding alternative data sources, which is challenging.

Researchers should consider using excess mortality to validate preliminary data, provided that reliable all-cause mortality statistics are available. Unlike other indicators requiring detailed mortality data, excess mortality can be determined more easily, as it only requires the total number of deaths, which is typically more available than detailed mortality data.

While it is easy to spot the discrepancy ex-post, how can a researcher know that the data is questionable ex-ante? Indicators such as CFR (which is part of the same databases) showed that Serbia had been a significant outlier within the Balkans since the pandemic began.

If researchers need COVID-19 mortality data with higher time resolution, we recommend using weekly excess mortality data (which is available for Serbia).

Editors and reviewers should also be cognizant of Serbia's COVID-19 preliminary data issue and examine how it may impact studies in which statistical analyses could be influenced by such an outlier.

Governments can benefit from prioritizing the development of robust systems and tools for releasing dependable preliminary vital statistics data. This approach can help prepare countries for future public health events that will benefit from ongoing and reliable monitoring of mortality, rather than having to create alternative and limited data streams on specific causes of death. The benefits of stable and robust vital statistics systems capable of producing provisional data for mortality surveillance have been demonstrated in many countries, such as the US and UK. Reducing the lag in the dissemination of vital statistics data ought to be an important goal for the governments in order to enable quality research needed for evidence-based decision making.

Limitation of our analysis

While vital statistics data is valuable for its accuracy and completeness, the pipeline for its generation is incredibly slow by today's standards. Our study's comparison was limited to 2020 and 2021, while omitting another deadly year (in COVID-19 terms), that of 2022.

It should be noted that our study only analysed a selection of prominent online databases that contained the problematic COVID-19 mortality data from Serbia. There are likely many more research studies, both published and unpublished, that utilize this data. As such, the full extent of the impact of this issue may be even greater than what our study was able to capture.

Our investigation focused solely on the COVID-19 mortality data and did not address other data produced daily by the same pipeline in Serbia, such as the number of persons tested, number of confirmed cases, or the number of vaccine shots administered.

CONCLUSIONS

The challenge of inconsistent COVID-19 mortality data, as addressed in this study, will continue to persist in the future. As the pandemic has yet to conclude, Serbia continues to generate flawed data. Consequently, research studies that rely on these data are also increasing in number.

Is it advisable for researchers to utilize preliminary COVID-19 data? For Serbia, the definitive answer is no. However, in a broader context, the use of such data may be acceptable, provided that authors, reviewers, and editors remain vigilant of any outliers and their potential to distort the findings.

What about other countries? Serbia merely serves as an example of a more pervasive issue that could arise elsewhere, not only in the context of this pandemic but also in potential future pandemics. Researchers ought to maintain vigilance regarding this problem and, when feasible, employ excess mortality data to corroborate any preliminary mortality data they might be using. In cases where researchers have reservations or concerns about the reliability of preliminary COVID-19 mortality data, considering excess mortality data as an alternative is advisable.

Database designers and managers should proactively incorporate quality control measures, cross-comparisons, and validation processes into their systems to ensure the integrity of the data being used by researchers, epidemiologists, and policy makers. Moreover, epidemiologists and policy makers should be mindful of potential data quality issues and exercise caution when interpreting and making decisions based on preliminary data, ensuring that policies and interventions are informed by reliable evidence.

Data statement

Data is available in a public, open access repository at https://osf.io/769me/?view_only=3c3c83167e144a399332ded846983c8e (anonymised for peer review)

REFERENCES

[1] Magdalenic I, Galjak M. Ageing map of the Balkan Peninsula. JGI JC SASA 2016;66:75–89. https://doi.org/10.2298/IJGI1601075M.
[2] Stosic S, Karanovic N. Health care economics in Serbia: Current problems and changes. VSP 2014;71:1055–61. https://doi.org/10.2298/VSP120205002S.
[3] Ministry of Health of the Republic of Serbia. Serbia's COVID-19 Website 2020. https://web.archive.org/web/20201230223540/https://covid19.rs/homepage-english/ (accessed October 9, 2021).
[4] Hannah Ritchie DB Edouard Mathieu, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Esteban Ortiz-Ospina, Joe Hasell, Bobbie Macdonald, Roser M. Coronavirus Pandemic (COVID-19). Our World in Data 2020.
[5] Marinković I, Galjak M. Excess mortality in Europe and Serbia during the COVID-19 pandemic in 2020. STNV 2021;59. https://doi.org/10.2298/STNV2101061M.
[6] Tozer J, González M. TheEconomist/covid-19-excess-deaths-tracker 2021.
[7] Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. ELife 2021;10:e69336. https://doi.org/10.7554/eLife.69336.
[8] Nurković JS. COVID-19 Death in Novi Pazar - "Serbian Bergamo." Disaster Med Public Health Prep 2021:1–2. https://doi.org/10.1017/dmp.2021.313.
[9] Government of Serbia. Conclusion on the establishment of unified and centralised software Information system Covid-19 (Zaključak Vlade o uspostavljanju jedinstvenog i centralizovanog softverskog rešenja — Informacioni sistem COVID-19 (IS COVID/19). 2020.
[10] Gordana Bjelobrk. Challenges in collecting data on mortality during the pandemic. STNV 2021;59.
[11] Republic of Serbia Statistical Office of the Republic of Serbia. Vital events, 2020. Statistical Release 2021;180.
[12] World Health Organization. WHO Mortality Database 2022. https://www.who.int/data/data-collection-tools/who-mortality-database (accessed August 29, 2022).
[13] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious Diseases 2020;20:533–4. https://doi.org/10.1016/S1473-3099(20)30120-1.
[14] Worldometers.info. Worldometer. Worldometer n.d. http://www.worldometers.info/ (accessed August 28, 2022).
[15] Harzing AW. Publish or Perish 2007.
[16] Corporation for Digital Scholarship. Zotero n.d.
[17] Kelly M, Mathenge G, Rao C. Lessons Learnt and Pathways forward for National Civil Registration and Vital Statistics Systems after the COVID-19 Pandemic: JEGH 2021;11:262. https://doi.org/10.2991/jegh.k.210531.001.
[18] CDC. COVID Data Tracker. Centers for Disease Control and Prevention 2020. https://covid.cdc.gov/covid-data-tracker (accessed March 15, 2023).

[19] Ahmad FB, Anderson RN, Knight K, Rossen LM, Sutton PD. Advancements in the National Vital Statistics System to Meet the Real-Time Data Needs of a Pandemic. Am J Public Health 2021;111:2133–40. https://doi.org/10.2105/AJPH.2021.306519.
[20] Escobar LE, Molina-Cruz A, Barillas-Mury C. BCG vaccine protection from severe coronavirus disease 2019 (COVID-19). Proc Natl Acad Sci USA 2020;117:17720–6. https://doi.org/10.1073/pnas.2008410117.
[21] Bârgăoanu A, Corbu N, Buturoiu R, Durach F. Managing the COVID-19 pandemic: predictors of trust in institutions in Romania. K 2021;ahead-of-print. https://doi.org/10.1108/K-12-2020-0913.
[22] Esaiasson P, Sohlberg J, Ghersetti M, Johansson B. How the coronavirus crisis affects citizen trust in institutions and in unknown others: Evidence from 'the Swedish experiment.' European Journal of Political Research 2021;60:748–60. https://doi.org/10.1111/1475-6765.12419.
[23] Corrao G, Rea F, Blangiardo GC. Lessons from COVID-19 mortality data across countries. Journal of Hypertension 2021;39:856–60. https://doi.org/10.1097/HJH.0000000000002833.
[24] Stevanović A, Šantrić-Milićević M. Modeling as an approach to pandemic uncertainty management: Mortality assessment of the COVID-19 pandemic. Srpski Medicinski Časopis Lekarske Komore 2021;2:278–85. https://doi.org/10.5937/smclk2-33796.
[25] Boytchev H. Covid-19: Why the Balkans' vaccine rollout lags behind most of Europe. BMJ 2021:n2412. https://doi.org/10.1136/bmj.n2412.
[26] Jovanović N. Korona: Broj umrlih i zaraženih višestruko veći od zvanično saopštenog. BIRN 2020.
[27] Raleigh VS. Tackling UK's mortality problem: covid-19 and other causes. BMJ 2020:m2295. https://doi.org/10.1136/bmj.m2295.
[28] Gosling T, Inotai E, Edward S, Claudia C. Democracy Digest: Hungary Vies with Czechia for World's Worst COVID Death Rate. Balkan Insight 2021. https://web.archive.org/web/20210530231914/https://balkaninsight.com/2021/03/26/democracy-digest-hungary-vies-with-czechia-for-worlds-worst-covid-death-rate/ (accessed May 30, 2021).
[29] Silva L, Figueiredo Filho D. Using Benford's law to assess the quality of COVID-19 register data in Brazil. Journal of Public Health 2021;43:107–10. https://doi.org/10.1093/pubmed/fdaa193.
[30] Galvêas D, Barros F, Fuzo CA. A forensic analysis of SARS-CoV-2 cases and COVID-19 mortality misreporting in the Brazilian population. Public Health 2021;196:114–6. https://doi.org/10.1016/j.puhe.2021.05.010.
[31] Kupek E. How many more? Under-reporting of the COVID-19 deaths in Brazil in 2020. Trop Med Int Health 2021;26:1019–28. https://doi.org/10.1111/tmi.13628.
[32] Biswas RK, Afiaz A, Huq S. Underreporting COVID-19: the curious case of the Indian subcontinent. Epidemiol Infect 2020;148:e207. https://doi.org/10.1017/S0950268820002095.
[33] Kilani A. Authoritarian regimes' propensity to manipulate Covid-19 data: a statistical analysis using Benford's Law. Commonwealth & Comparative Politics 2021;59:319–33. https://doi.org/10.1080/14662043.2021.1916207.
[34] Caporali A, Garcia J, Couppié É, Poniakina S, Barbieri M, Bonnet F, et al. The demography of COVID-19 deaths database, a gateway to well-documented international data. Sci Data 2022;9:93. https://doi.org/10.1038/s41597-022-01191-y.

Appendices

Appendix A
Detailed search strategy

We used Web of Science (WoS) to search for references to the three databases. After identifying the articles citing these databases, we refined the search results to those mentioning mortality and those with names of countries and the word "global" in the title, keywords, or appendix, or those explicitly mentioning Serbia. That way, we included articles that involve cross-country comparisons (that could have potentially been exposed to the erroneous data coming from Serbia), and those that have something to do with Serbia.

We used Google Scholar (GS) to search for references to the Serbian government's sites. GS was used to expand the search beyond the WoS, as many Serbian journals are not indexed in the WoS. GS search also enabled us to find the research published in Serbian Language only.

Search string for WoS search:

Citations of ourworldindata (Cited Work) or **owid** (Cited Work) or **Our World in Data** (Cited Work) or **worldometers** (Cited Work) or **worldometers.info** (Cited Work) or **10.1016/S1473-3099(20)30120-1** (Cited DOI) or **An interactive web-based dashboard to track COVID-19 in real time** (Cited Work) or **COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University** (Cited Work) and **2020-2022** (Cited Year(s)) and **Africa** (Exclude – Search within all fields) and **Asia** (Exclude – Search within all fields) and **Deaths Or Mortality** (Search within all fields) and **"Excess Mortality"** (Exclude – Search within all fields) and **Countries Or Global** (Search within all fields)

Search string for GS search (performed through publish or perish software):

("**covid19.data.gov.rs**" OR "**covid19.rs**") AND ("**mortality**" OR "**deaths**" OR "**smrti**" OR "**umrli**")

Appendix B

Title	Authors	Year	Publication	DOI
Escobar, LE. et al.	BCG vaccine protection from severe coronavirus disease 2019 (COVID-19)	2020	Proc. Natl. Acad. Sci. U. S. A.	10.1073/pnas.2008410117
Lal, P. et al.	The dark cloud with a silver lining: Assessing the impact of the SARS COVID-19 pandemic on the global environment	2020	Sci. Total Environ.	10.1016/j.scitotenv.2020.139297
Guo, C. et al.	Meteorological factors and COVID-19 incidence in 190 countries: An observational study	2021	Sci. Total Environ.	10.1016/j.scitotenv.2020.143783
Dickens, BL. et al.	Strategies at points of entry to reduce importation risk of COVID-19 cases and reopen travel	2020	J. Travel Med.	10.1093/jtm/taaa141
Blasius, B	Power-law distribution in the number of confirmed COVID-19 cases<?A3B2 show [editpick]?>	2020	Chaos	10.1063/5.0013031

Shapira, G. et al.	Ethnic differences in alpha-1 antitrypsin deficiency allele frequencies may partially explain national differences in COVID-19 fatality rates	2020	Faseb J.	10.1096/fj.202002097
Hale, T. et al.	Government responses and COVID-19 deaths: Global evidence across multiple pandemic waves	2021	PLoS One	10.1371/journal.pone.0253116
Karic, T; Mededovic, J	Covid-19 conspiracy beliefs and containment-related behaviour: The role of political trust	2021	Pers. Individ. Differ.	10.1016/j.paid.2021.110697
Coccia, M.	Optimal levels of vaccination to reduce COVID-19 infected individuals and deaths: A global analysis	2022	Environ. Res.	10.1016/j.envres.2021.112314
Petropoulos, F. et al.	COVID-19: Forecasting confirmed cases and deaths with a simple time series model	2022	Int. J. Forecast.	10.1016/j.ijforecast.2020.11.010
Hogan, AB. et al.	Within-country age-based prioritisation, global allocation, and public health impact of a vaccine against SARS-CoV-2: A mathematical modelling analysis	2021	Vaccine	10.1016/j.vaccine.2021.04.002
Colson, P. et al.	Analysis of SARS-CoV-2 Variants From 24,181 Patients Exemplifies the Role of Globalization and Zoonosis in Pandemics	2022	Front. Microbiol.	10.3389/fmicb.2021.786233
Kolarevic, S. et al.	Detection of SARS-CoV-2 RNA in the Danube River in Serbia associated with the discharge of untreated wastewaters	2021	Sci. Total Environ.	10.1016/j.scitotenv.2021.146967
Post, L et al.	Surveillance of the Second Wave of COVID-19 in Europe: Longitudinal Trend Analyses	2021	JMIR Public Health Surveill.	10.2196/25695
Vyklyuk, Y et al.	Modeling and analysis of different scenarios for the spread of COVID-19 by using the modified multi-agent systems - Evidence from the selected countries	2021	Results Phys.	10.1016/j.rinp.2020.103662
Erman, A; Medeiros, M	Exploring the Effect of Collective Cultural Attributes on Covid-19-Related Public Health Outcomes	2021	Front. Psychol.	10.3389/fpsyg.2021.627669
Bretschger, L et al.	COVID-19 infections and fatalities developments: empirical evidence for OECD countries and newly industrialized economies	2020	Int. Econ. Econ. Policy	10.1007/s10368-020-00487-x
Cepaluni, G; Dorsch, MT; Branyiczki, R	Political regimes and deaths in the early stages of the COVID-19 pandemic	2022	J. Public Financ. Public Choice	10.1332/251569121X16268740317724
Zhai, YH; Jiang, DY; Gozgor, G; Cho, EH	The Amplifying Effect of Conflicts on Case Fatality Rate of COVID-19: Evidence From 120 Countries	2021	Front. Public Health	10.3389/fpubh.2021.681604
Pincombe, M; Reese, V; Dolan, CB	The effectiveness of national-level containment and closure policies across income levels during the COVID-19 pandemic: an analysis of 113 countries	2021	Health Policy Plan.	10.1093/heapol/czab054
Rajkumar, RP.	Cross-National Variations in COVID-19 Mortality: The Role of Diet, Obesity and Depression	2021	Diseases	10.3390/diseases9020036
Yu, CS. et al.	A COVID-19 Pandemic Artificial Intelligence-Based System With Deep Learning Forecasting and Automatic Statistical Data Acquisition: Development and Implementation Study	2021	J. Med. Internet Res.	10.2196/27806
Lenton, TM; Boulton, CA; Scheffer, M	Resilience of countries to COVID-19 correlated with trust	2022	Sci Rep	10.1038/s41598-021-03358-w
Ngepah, N	Socio-economic determinants of global COVID-19 mortalities: policy lessons for current and future pandemics	2021	Health Policy Plan.	10.1093/heapol/czaa161

Jagodic, J. et al.	Possible zinc deficiency in the Serbian population: examination of body fluids, whole blood and solid tissues	2021	Environ. Sci. Pollut. Res.	10.1007/s11356-021-14013-2
Jovicic-Bata, J. et al.	Coping with the burden of the COVID-19 pandemic: a cross-sectional study of community pharmacists from Serbia	2021	BMC Health Serv. Res.	10.1186/s12913-021-06327-1
Singh, J. et al.	Role of multiple factors likely contributing to severity-mortality of COVID-19	2021	Infect. Genet. Evol.	10.1016/j.meegid.2021.105101
Strizovic, S. et al.	Influence of COVID-19 pandemic on quality of life in patients with epilepsy-Follow-up study	2021	Epilepsy Behav.	10.1016/j.yebeh.2021.108026
Ando, M; Hayakawa, K	Impact of COVID-19 on trade in services	2022	Jpn. World Econ.	10.1016/j.japwor.2022.101131
Phannajit, J et al.	Factors Associated with the Incidence and Mortality of Coronavirus Disease 2019 (COVID-19) after 126-million Cases: A Meta-analysis	2021	J. Epidemiol. Glob. Health	10.2991/jegh.k.210527.001
Quilodran, CS. et al.	Air temperature influences early Covid-19 outbreak as indicated by worldwide mortality	2021	Sci. Total Environ.	10.1016/j.scitotenv.2021.148312
Chatterjee, B; Karandikar, RL; Mande, SC	Mortality due to COVID-19 in different countries is associated with their demographic character and prevalence of autoimmunity	2021	Curr. Sci.	10.18520/cs/v120/i3/501-508
Marziali, M. et al.	Predictors of COVID-19 testing rates: A cross-country comparison	2021	Int. J. Infect. Dis.	10.1016/j.ijid.2020.12.083
Stanojevic, S. et al.	Simulation and prediction of spread of COVID-19 in The Republic of Serbia by SEAIHRDS model of disease transmission	2021	Microb. Risk Anal.	10.1016/j.mran.2021.100161
Docquier, F. et al.	Cross-border mobility responses to COVID-19 in Europe: new evidence from facebook data	2022	Global. Health	10.1186/s12992-022-00832-6
Dempere, J.	A recipe to control the first wave of COVID-19: More or less democracy?	2021	Transform. Gov.-People Process Policy	10.1108/TG-08-2020-0206
Kosic, A. et al.	A Cross-Cultural Study of Distress during COVID-19 Pandemic: Some Protective and Risk Factors	2021	Int. J. Environ. Res. Public Health	10.3390/ijerph18147261
Heuveline, P	Global and National Declines in Life Expectancy: An End-of-2021 Assessment	2022	Popul. Dev. Rev.	10.1111/padr.12477
de Oliveira, AMB et al.	Using GAM functions and Markov-Switching models in an evaluation framework to assess countries' performance in controlling the COVID-19 pandemic	2021	BMC Public Health	10.1186/s12889-021-11891-6
Rajkumar, RP	Is There a Relationship Between ADHD and COVID-19 Prevalence and Mortality Indices? An Analysis of Data From 156 Countries	2022	J. Atten. Disord.	10.1177/10870547211056894
Korneta, P. et al.	Mutual relationships between SARS-CoV-2 test numbers, fatality and morbidity rates	2021	BMC Public Health	10.1186/s12889-021-12021-y
Morshed, MM. et al.	Common factors of COVID-19 cases and deaths among the most affected 50 countries	2021	Diabetes Metab. Syndr.-Clin. Res. Rev.	10.1016/j.dsx.2021.102247
Thomas, BS. et al.	Estimating the case fatality ratio for COVID-19 using a time-shifted distribution analysis	2021	Epidemiol. Infect.	10.1017/S0950268821001436
Banski, J. et al.	Socioeconomic Conditioning of the Development of the COVID-19 Pandemic and Its Global Spatial Differentiation	2021	Int. J. Environ. Res. Public Health	10.3390/ijerph18094802
Ho, YH. et al.	COVID-19 Pandemic Analysis for a Country's Ability to Control the Outbreak Using Little's Law: Infodemiology Approach	2021	Sustainability	10.3390/su13105628
Cvetkovic, VM. et al.	A Predictive Model of Pandemic Disaster Fear Caused by Coronavirus (COVID-19): Implications for Decision-Makers	2022	Int. J. Environ. Res. Public Health	10.3390/ijerph19020652
Lin, BB. et al.	Revealing the linguistic and geographical disparities of public awareness to Covid-19 outbreak through social media	2022	Int. J. Digit. Earth	10.1080/17538947.2022.2070677

Khafaga, DS. et al.	Intelligent Model for Data Analytical Study of Coronavirus COVID-19 Databases	2022	Electronics	10.3390/electronics11131975
Su, DJ; Alshehri, K; Pagan, JA	Income inequality and the disease burden of COVID-19: Survival analysis of data from 74 countries	2022	Prev. Med. Rep.	10.1016/j.pmedr.2022.101828
Wahltinez, O. et al.	COVID-19 Open-Data a global-scale spatially granular meta-dataset for coronavirus disease	2022	Sci. Data	10.1038/s41597-022-01263-z
Pourmalek, F	CovidVisualized: Visualized compilation of international updated models' estimates of COVID-19 pandemic at global and country levels	2022	BMC Res. Notes	10.1186/s13104-022-06020-4
Erdem, S. et al.	Investigating the effect of macro-scale estimators on worldwide COVID-19 occurrence and mortality through regression analysis using online country-based data sources	2022	BMJ Open	10.1136/bmjopen-2021-055562
Wang, LY. et al.	Vaccination associated with gross domestic product and fewer deaths in countries and regions A verification study	2022	Medicine (Baltimore)	10.1097/MD.0000000000028619
Moozhipurath, RK; Kraft, L	Association of lockdowns with the protective role of ultraviolet-B (UVB) radiation in reducing COVID-19 deaths	2021	Sci Rep	10.1038/s41598-021-01908-w
Rajkumar, RP	The relationship between pre-COVID prevalence of common mental disorders and the impact of COVID-19	2021	Minerva Psychiat.	10.23736/S2724-6612.20.02098-1
Zhu, J; Gallego, B	Evolution of disease transmission during the COVID-19 pandemic: patterns and determinants	2021	Sci Rep	10.1038/s41598-021-90347-8
Feng, X. et al.	On fluctuating characteristics of global COVID-19 cases and identification of inflection points	2021	Libr. Hi Tech	10.1108/LHT-10-2020-0263
Sajic, JL. et al.	System identification and mathematical modeling of the pandemic spread COVID-19 in Serbia	2022	IFAC PAPERSONLINE	10.1016/j.ifacol.2022.06.003
Nurkovic, JS	COVID-19 Death in Novi Pazar-Serbian Bergamo		Dis. Med. Public Health Prep.	10.1017/dmp.2021.313
Jovanovic, R. et al.	Modelling Voluntary General Population Vaccination Strategies during COVID-19 Outbreak: Influence of Disease Prevalence	2021	Int. J. Environ. Res. Public Health	10.3390/ijerph18126217

Table B.1 Studies that include mortality data from Serbia.

Title	**Authors**	**Year**	**Publication**	**DOI**	**No. Countries**
On fluctuating characteristics of global COVID-19 cases and identification of inflection points	Feng, Xin; Zhang, Hanshui; Zhang, Yue; Sun, Liming; Li, Jiapei; Wu, Ye	2021	Library Hi Tech	10.1108/LHT-10-2020-0263	106
Mutual relationships between SARS-CoV-2 test numbers, fatality and morbidity rates	Korneta, Piotr; Zawiła-Niedźwiecki, Janusz; Domański, Jarosław	2021	BMC Public Health	10.1186/s12889-021-12021-y	156
The effectiveness of national-level containment and closure policies across income levels during the COVID-19 pandemic: an analysis of 113 countries	Pincombe, Morgan; Reese, Victoria; Dolan, Carrie B	2021	Health Policy and Planning	10.1093/heapol/czab054	107

A recipe to control the first wave of COVID-19: more or less democracy?	Dempere, Juan	2021	Transforming Government: People, Process and Policy	10.1108/TG-08-2020-0206	94
Predictors of COVID-19 testing rates: A cross-country comparison	Marziali, Megan E.; Hogg, Robert S.; Oduwole, Oluwamayowa A.; Card, Kiffer G.	2021	International Journal of Infectious Diseases	10.1016/j.ijid.2020.12.083	156
The relationship between pre-COVID prevalence of common mental disorders and the impact of COVID-19	Rajkumar, Ravi P.	2021	Minerva Psychiatry	10.23736/S2724-6612.20.02098-1	NA
Socio-economic determinants of global COVID-19 mortalities: policy lessons for current and future pandemics	Ngepah, Nicholas	2021	Health Policy and Planning	10.1093/heapol/czaa161	113
Political regimes and deaths in the early stages of the COVID-19 pandemic	Cepaluni, Gabriel; Dorsch, Michael T.; Branyiczki, Réka	2022	Journal of Public Finance and Public Choice	10.1332/251569121X16268740317724	84
Is There a Relationship Between ADHD and COVID-19 Prevalence and Mortality Indices? An Analysis of Data From 156 Countries	Rajkumar, Ravi Philip	2022	Journal of Attention Disorders	10.1177/10870547211056894	156
Cross-National Variations in COVID-19 Mortality: The Role of Diet, Obesity and Depression	Rajkumar, Ravi Philip	2021	Diseases	10.3390/diseases9020036	196

Table B.2 Studies that likely include mortality data from Serbia.

Appendix D

Country	Preliminary	Final	Absolute Difference	Ratio (Final/Preliminary)
Austria	7486	6479	-1007	0.87
Bosnia and Herzegovina	4050	4438	388	1.10
Bulgaria	7576	8568	992	1.13
Croatia	3920	4506	586	1.15
Cyprus	120	131	11	1.09
Czechia	11580	10570	-1010	0.91
Estonia	229	203	-26	0.89
Germany	33071	39886	6815	1.21
Hungary	9537	8979	-558	0.94
Iceland	29	30	1	1.03
Latvia	635	701	66	1.10
Lithuania	1800	2256	456	1.25
Netherlands	11459	20212	8753	1.76

North Macedonia	2503	2810	307	1.12
Poland	28554	41486	12932	1.45
Serbia	3211	10356	7145	3.23
Slovakia	2138	3726	1588	1.74
Slovenia	2697	3399	702	1.26
Spain	50837	74839	24002	1.47

Table D.1 Differences between the preliminary (daily reported) and final (vital statistics) number of COVID-19 deaths in European countries with available vital statistics data for 2020.

Source: WHO and Eurostat