Excess Mortality and its Determinants During the COVID-19 Pandemic in 21 Countries: An Ecological Study from the C-MOR Project, 2020 and 2021

Journal of Epidemiology and Global Health
https://doi.org/10.1007/s44197-024-00320-7

RESEARCH ARTICLE

Excess Mortality and its Determinants During the COVID-19 Pandemic in 21 Countries: An Ecological Study from the C-MOR Project, 2020 and 2021

Mohammad Reza Rahmanian Haghighi[1,2] · **Chryso Th Pallari**[1] · **Souzana Achilleos**[1] · **Annalisa Quattrocchi**[1] · **John Gabel**[3] · **Andreas Artemiou**[4] · **Maria Athanasiadou**[5] · **Stefania Papatheodorou**[6,7] · **Tianyu Liu**[8] · **José Antonio Cernuda Martínez**[9] · **Gleb Denissov**[10] · **Błażej Łyszczarz**[11] · **Qian Huang**[12] · **Kostas Athanasakis**[13] · **Catherine M. Bennett**[14] · **Claudia Zimmermann**[15] · **Wenjing Tao**[16] · **Serge Nganda Mekogo**[17] · **Terje P. Hagen**[18] · **Nolwenn Le Meur**[19] · **Jackeline Christiane Pinto Lobato**[20] · **Giuseppe Ambrosio**[21] · **Ivan Erzen**[22] · **Binyamin Binyaminy**[23] · **Julia A. Critchley**[24] · **Lucy P. Goldsmith**[25] · **Olesia Verstiuk**[26] · **Jideofor Thomas Ogbu**[27] · **Laust H. Mortensen**[28] · **Levan Kandelaki**[29] · **Marcin Czech**[30] · **Joseph Cutherbertson**[31] · **Eva Schernhammer**[32] · **Catharina Vernemmen**[17] · **Antonio José Leal Costa**[33] · **Tamar Maor**[34] · **Dimos Alekkou**[35] · **Bo Burström**[16] · **Antonis Polemitis**[36] · **Andreas Charalambous**[37] · **Christiana A. Demetriou**[1]

Received: 28 June 2024 / Accepted: 19 October 2024

Abstract

Introduction The COVID-19 pandemic overwhelmed health systems, resulting in a surge in excess deaths. This study clustered countries based on excess mortality to understand their response to the pandemic and the influence of various factors on excess mortality within each cluster.

Materials and Methods This ecological study is part of the COVID-19 MORtality (C-MOR) Consortium. Mortality data were gathered from 21 countries and were previously used to calculate weekly all-cause excess mortality. Thirty exposure variables were considered in five categories as factors potentially associated with excess mortality: population factors, health care resources, socioeconomic factors, air pollution, and COVID-19 policy. Estimation of Latent Class Linear Mixed Model (LCMM) was used to cluster countries based on response trajectory and Generalized Linear Mixture Model (GLMM) for each cluster was run separately.

Results Using LCMM, two clusters were reached. Among 21 countries, Brazil, the USA, Georgia, and Poland were assigned to a separate cluster, with the mean of excess mortality z-score in 2020 and 2021 around 4.4, compared to 1.5 for all other countries assigned to the second cluster. In both clusters the population incidence of COVID-19 had the greatest positive relationship with excess mortality while interactions between the incidence of COVID-19, fully vaccinated people, and stringency index were negatively associated with excess mortality. Moreover, governmental variables (government revenue and government effectiveness) were the most protective against excess mortality.

Conclusion This study highlighted that clustering countries based on excess mortality can provide insights to gain a broader understanding of countries' responses to the pandemic and their effectiveness.

Keywords COVID-19 · Excess mortality · Public health measures · Vaccination rate · Governance

1 Introduction

The reported number of COVID-19 deaths during the pandemic may not reflect the true extent of the health burden or the total lives lost as a result of the pandemic. Excess mortality is defined by the WHO as "the mortality above what would be expected based on the non-crisis mortality rate in the population of interest" [1]. By considering 74 nations, COVID-19 Excess Mortality Collaborators (2022) estimated that although the globally reported COVID-19 deaths were 5.94 million between January 1, 2020, and December 31, 2021, the projected global death toll from the COVID-19 pandemic was 18.2 million [2]. Similarly,

Extended author information available on the last page of the article

Published online: 11 November 2024

Msemburi et al. (2022) utilized the Bayesian Poisson framework to demonstrate that, in contrast to the 5.42 million deaths attributed to COVID-19 reported to the WHO in 2020 and 2021, the estimated global excess mortality is 2.74 times higher, at 14.83 million [3].

Previous studies have evaluated various factors as potential determinants of excess mortality during the COVID-19 pandemic. A study in England during the first wave of the COVID-19 pandemic, considering only people aged 40 or older, found that communities with a higher risk of excess mortality had a high density of care facilities and/or a large proportion of residents on financial assistance, living in overcrowded houses, and of non-white ethnicity [4]. Furthermore, by considering 79 high-, middle- and low-income countries and using median quantile regression, Kapitsinis (2021) found that privatization of the health care sector, inadequate public health spending, and delayed adoption of preventative measures all contributed significantly to excess mortality during the pandemic, whereas health expenditure and the number of hospital beds and doctors per 100,000 population were negatively associated with excess mortality. The study concluded that an adequately funded healthcare system, aligned with universal access and robust primary healthcare, could be a good measure to control excess mortality during the pandemic [5]. Another study of 213 countries found that strong social cohesion and effective risk communication, including factors such as trust in government and public programs, were associated with fewer excess deaths from COVID-19 [6].However, an increased number of COVID-19-related disorders and a lack of government transparency have been associated with higher death rates. Furthermore, countries with higher unemployment rates reported higher excess mortality [6]. Lastly, Sun et al. (2023) investigated 80 countries and showed that pre-existing conditions such as age and health risks, as well as public trust in healthcare systems, play important roles in explaining why death rates differ across countries [7]. Besides these factors, immunization against COVID-19 has significantly altered the path of the pandemic by saving tens of millions of lives around the world [8]. A cross-sectional study among 173 countries found that higher vaccination coverage was strongly associated with lower overall mortality, highlighting the critical role of vaccination in reducing deaths during the pandemic [9]. In addition, non-pharmaceutical measures, particularly lockdowns and social distancing, have been shown to significantly decrease the overall transmission rate in 2020, not just in the immediate intervention period [10]. Interestingly, however, at the ecological level, stringency index of governmental control measures was associated with higher excess mortality in 2020 but lower mortality in 2021 [11]. Furthermore, a recent cross-national study showed that weekly COVID-19 incidence was significantly associated with excess mortality in both years, however, this relationship weakened in 2021 as vaccination rates increased [11].

Though several studies have previously investigated the characteristics associated with increased mortality rate during COVID-19, they mostly considered pre-defined country income categories [12, 13] or only focused on one geographic region [14–16]. The purpose of this study was to evaluate the feasibility of clustering countries from different world regions based on their excess mortality experience to gain a better understanding of their response to the pandemic. This would in turn enable investigation of the relationship of various factors on each cluster's excess mortality during the COVID-19 pandemic. Twenty-one different countries and/or regions that are a part of the COVID-19 MORtality (C-MOR) Consortium were investigated during 2020 and 2021. This study examines how variations in government policies and socio-economic factors across countries impacted excess mortality during the COVID-19 pandemic. The theoretical model being tested with latent variable analysis hypothesizes that various factors contributing to excess mortality will be found in common across countries with similar excess death profiles. Latent variable analysis is applied to understand what distinguishes these clusters, providing insight into the potential drivers of excess death outcomes in pandemics. Thus, this study used a data-driven approach to cluster countries based on their excess mortality patterns. Determinants of excess mortality were then evaluated within each cluster, and the factors contributing to differences between clusters were identified.

2 Data Analysis and Methods

Twenty-one countries and/or regions participating in the worldwide consortium, provided mortality data that were analyzed in this ecological study. Countries that were included in this study are Australia, Austria, Belgium, Brazil, Cyprus, Denmark, England and Wales, Estonia, France, Georgia, Greece, Israel, Italy, Northern Ireland, Norway, Poland, Slovenia, Spain, Sweden, Ukraine, and the United States of America (USA). Additional countries in the consortium that provided data but had a completeness of vital registration systems of < 90%, were excluded from this investigation (Mauritius, Peru, and Kazakhstan) to enhance the reliability and accuracy of the analysis. This investigation builds on the work of a preceding study, where data on all-cause deaths by age and gender were gathered from national vital statistics databases every week from 2015 through 2021 [11]. The methodology used to estimate excess all-cause mortality was described previously [8]. Briefly, crude mortality rates (CMRs) were first calculated for the total population using population estimates from the World Bank (World Bank Group, 2023), the UK Office for National Statistics (Office for National Statistics, 2024) and Eurostat (Eurostat, 2023). Age-specific mortality rates (ASpMRs) and weekly (directly) age-standardised mortality rates (ASMRs) were then calculated using the aggregated age groups provided by each country, as detailed in that publication. Then, total weekly excess mortality for 2020 and 2021 was calculated by comparing the observed weekly ASMR (per 100 000 population) against the expected weekly ASMR, estimated based on a time series regression analysis of historical data (2015– 2019) as previously described [11, 17–21]The regression models were built on complete weeks; truncated weeks were excluded. Finally, the z-score (number of observed deaths—expected mortality)/standard deviation of the residuals) of these data was calculated, and a z-score of more than four was considered as substantial excess mortality [17].

The current study used the weekly excess mortality z-score for 2020–2021 as the outcome variable. Five main categories (comprising a total of thirty potential predictor variables) were considered for evaluating their impact on excess mortality: population factors, health resources, socioeconomic factors, air pollution, and COVID-19 policy.

Country-level population factors consist of the population size and density, median age, percent of people over 65 years of age, median age, life expectancy, and prevalence of hypertension, diabetes, and obesity. From the health care resources aspect, hospital beds per thousand, the densities of nursing personnel and medical doctors (both per 10,000 population), universal health coverage (UHC), health expenditure as a share of Gross Domestic Product (GDP), total vaccinations per 100 population, fully vaccinated per 100 population, completeness of vital registration, healthcare access, and quality index (HAQ), and COVID incidence per 100 population have been considered. Socioeconomic status consists of GDP, Human Development Index (HDI), unemployment rate, government revenue, prosperity index, Gini index (measures income inequality), control of corruption, government effectiveness, Inequality-adjusted Human Development Index (IHDI), Gender Inequality Index (GII). From the COVID-19 policy perspective, a Government Response stringency index (composite measure based on nine response indicators including school closures, workplace closures, and travel [22]) was investigated, and finally, we used fine particulate matter (PM2.5) as a variable reflecting air pollution. Supplementary Table S 1 provides the source for each of the aforementioned potential predictor variables as well as an explanation of each variable and the range of possible values. Briefly, data were retrieved from various sources including the World Bank, World Health Organization (WHO) and Our World in Data. Among these variables, COVID incidence, vaccination rates, and stringency index are time-variant predictors, changing on a weekly basis. In these analyses, 21 countries/regions were evaluated. Each country or region was observed for 90 to 104 weeks between 2020 and 2021. Two datasets were used: 1) To find the trajectory of excess mortality, a database with 2172 observations (weekly data for each country) was used; 2) For evaluating the relationship between predictors and the outcome, a dataset with 2000 observations was used (in this dataset, missing values existed, as data for COVID-19 incidence before January 2020 were not available). The models employed in this study are relatively robust to potential violations of assumptions. Decision trees can manage categorical features directly, without the need for preprocessing [23]. Mixed-effects models are sufficiently robust for researchers to use them even when distributional assumptions are not fully satisfied. Nonetheless, this does not exempt researchers from the necessity of rigorously assessing the model's accuracy [18]. Hence, a data engineering approach was applied to the dataset before running the data analysis. First, the Yeo-Johnson transformation was applied to z-scores based on outcome distribution [24]. After applying the Yeo-Johnson transformation, the model did not have obvious outliers. Scaling predictor variables (Min–max scaling between 0 and 1) was another data engineering feature applied to the dataset. Finally, to handle collinearity and avoid overfitting, the Variance Inflation Factor (VIF) function was used for feature selection. VIF less than 10 was considered a threshold [25], and features with VIF greater than 10 were removed sequentially. Data analyses were performed using RStudio (R version 4.3.1).

Given the multi-level nature of our datasets we required a sophisticated and robust clustering method like the Latent Class Analysis rather than the traditional k-means technique. Latent class analysis is a flexible statistical technique used to identify distinct groups or segments of data. This method has become a standard tool for identifying hidden patterns in various datasets and providing meaningful insights based on observed indicators [26]. Latent Class Analysis models provide a more robust statistical foundation than K-means and Hierarchical Clustering for both exploratory research and theory testing [27]. Additionally, the complexity of multi-level dataset requires a more sophisticated clustering method than traditional techniques like K-means. Countries were first classified according to their outcome trajectory (z-score trajectory) using the "Estimation of Latent Class Linear Mixed Model" (LCMM). Next, a Generalized Linear Mixture Model (GLMM) was run for each cluster separately to investigate all thirty predictor variables against z-score values as the outcome variable. Then, a Classification And Regression Tree (CART) algorithm was used to understand how time-variant and time-invariant predictors classified observations within their assigned clusters, across different models. It should be noted that CART analysis considers each observation separately, independent of country (like in a cross-sectional study); thus, we need to be conservative when interpreting the time-variant model outputs. For time-invariant predictors, because of the limited number of observations (21 observations), univariate logistic regression was used for comparing clusters. Details for each statistical analysis have been provided in Table 1.

3 Results

Variable trajectories in excess mortality were observed across the 21 studies investigated, with evidence for some countries, such as United States and Brazil, being more severely affected than others, such as Australia and Denmark.

Figure 1 illustrates the trajectory of z-scores for each country/region, where each point represents the six-month mean of the z-scores. The trajectory figure of the z-score for all data points is available in the Supplementary, Fig. S1. In addition, Supplementary Table S2 contains the descriptive analysis of kept variables after screening for collinearity.

Among various models for LCMM, a model with two latent classes showed better performance (lower Akaike information criterion (AIC) and Bayesian information criterion (BIC)) than other models and was thus retained. The full report of LCMM's outcome is presented in Supplementary Table S3. After running posterior probabilities, countries with homogenous trajectories were classified in the same class. The result shows the mean of excess mortality in the first cluster is about one-third of the second cluster. Table 2 presents the result of countries classification for the time-variant dataset. In addition, Fig. S1 shows the trajectory of z-score for each country/region.

After clustering, a GLMM mixed model, with country as a random effect and all other predictors as fixed effects, was run for each cluster separately. The first cluster consists of 17 countries with 1618 observations. The model's marginal and conditional R-squares were 0.217 and 0.914, respectively. A significantly higher value of the conditional R-squares compared to the marginal R-squared shows that a substantial percentage of the variation is explained by both the fixed effects and the random. The multilevel models employed to examine pandemic-related variables for each cluster reveal that the random effect of nation accounts for a significant percentage of the variation in the outcome. This variation is most notable for the first cluster (with lower excess mortality), which, by considering country-specific variance, conditional R-square (for random effect) is 70 percent points higher than marginal R-square (fix-effect). The coefficient of variables with a $p \leq 0.05$ is shown in Table 3 (the complete table is available in Supplementary Table S4). Among significant variables, "COVID incidence with three weeks lag" (β: 0.86, $p < 0.001$) exhibited the strongest relationship with the increase in excess mortality. On the other hand, among significant variables that were negatively associated with excess mortality, a higher rate of obesity (β: -1.19, $p < 0.05$), government revenue (β: − 0.7, $p < 0.001$) and elderly population (β: − 0.7, $p < 0.01$) have the strongest effects. In addition, the interaction between three time-variant variables (proportion of fully vaccinated with three weeks lag & COVID incidence with three weeks lag (β: − 0.53, $p < 0.01$), Stringency index & COVID incidence with three weeks lag (β: − 0.78, p = > 0.0505) and Stringency index & proportion of fully vaccinated with three weeks lag (β: − 0.11, $p < 0.01$)) had a negative relationship with excess mortality during the study period.

The second cluster only contains four countries and 382 observations. The marginal R-squared was 0.66, indicating that the model's independent variables explain 0.66 of the dependent variable's variance. Additionally, the conditional R-squared value is 0.69, representing the proportion of variance in the dependent variable explained by the independent variables. Among significant time-variant variables, similar to the first cluster, the incidence of COVID with three-weeks-lag had the highest positive relationship with the increase in excess mortality (β:8.22, $p < 0.01$) followed by elderly population rate (β: 1.9, $p < 0.01$). In addition, interaction variables are negatively associated with excess mortality, proportion of fully vaccinated with three weeks lag & COVID incidence with three weeks lag (β: -2.43, $p < 0.01$), and Stringency index & COVID incidence with three weeks lag (β: -6.37, $p < 0.01$). In contrast to the first cluster, a higher rate of population over 65 years old was related to higher excess mortality (β: 1.93, $p < 0.01$). The coefficient of variables with a p-value less than 0.1 is demonstrated in Table 4 (the complete table is available in Supplementary Table S 5).

To compare clusters based on their time-variant predictors, a CART model with tenfold cross-validation was used to find the optimal model. Figure 2 displays the CART plotted model. The top number in the node box is the CART model's predicted cluster. The middle two numbers show the number of observations from each cluster in that node (left for cluster one and right for the second cluster), and the bottom number indicates the percentage of observations. The outcome consists of five terminal nodes and four non-terminal nodes.

Variable importance and terminal nodes highlighted the importance of COVID-19 incidence. The classification model assigned observations with a low level of COVID incidence per 1000 (lower than 0.93) to cluster one (with lower excess mortality), which had a lower mean of z-score (left branch). Moreover, by considering the right branch, most observations with higher levels of proportion of fully vaccinated were assigned to the first cluster.

Due to the limited number of observations, our initial exploration with individual CART models for each variable

Table 1 Summary of statistical analyses used in this study

Analysis / Model	Estimation of Latent Class Linear Mixed Model (LCMM)
Purpose of using model	Clustering countries based on their trajectory
Description	Latent class mixed models explore varied population trajectories' latent profiles. To model trajectories, they use mixed models' theory to account for individual correlation in repeated measures and latent class models to distinguish homogenous latent clusters
Outcome	Z-score (Time-variant)
Predictors	Not applicable
Command/Package	"hmle" command from the "lcmm" package in R*
Analysis / Model	**Generalized Linear Mixture Model (GLMM)**
Purpose of using model	To assess fix and random effect of predictors on the outcome for each cluster
Description	Builds upon generalized linear models by incorporating random effects to account for correlated or clustered data. The significance level for this model is set at 0.05.
Outcome	Z-score (Time-variant)
Predictors	Three time-variant & their interactions were used for GLMM models. The number of time-invariants entered the GLMM analysis for each cluster was different. All non-colinear time-invariant predictors were entered into the GLMM model for the first cluster with 1618 observations. For the GLMM model for the second cluster with 382 observations, among time-invariant predictors, the number of hospital beds and health expenditure from the health resource category, population over 65 years old from the population category, government effectiveness, government revenue, unemployment, and Gini index were entered into the model Random effect: Country
Command/Package	"lmer" command from the "lme4" package in R**
Analysis / Model	**Classification And Regression Trees (CART)**
Purpose of using model	LCMM-derived clusters were compared using a CART model to analyze predictor relationships
Description	Basic decision trees divide the dataset into homogenous subgroups with comparable response values and fit a simple constant in each subgroup. Recursive binary partitions are used to construct subgroups (nodes) by asking yes/no questions about each feature. If the answer is "yes", the left branch will be considered, and vice versa. This process is repeated until a stopping point, such as a tree's maximum depth***. CART functions through recursive data partitioning. Recursive partitioning is a process that sequentially divides data into ever-smaller parts. It is "stagewise," not "stepwise," because earlier stages are not revisited once the outcomes of later stages are known****. Various metrics such as accuracy of predictions, sensitivity, specificity, precision, and F1 score were used to evaluate the performance of CART models
Outcome	The two clusters resulting from LCMM analysis: 0 for the first cluster & and 1 for the second cluster (binary outcome)
Predictors	Time variant predictors
Command/Package	"rpart" command from the "rpart" package in R*****
Analysis / Model	**Univariate Logistic Regression**
Purpose of using model	LCMM-derived clusters were compared using a Univariate Logistic Regression model to analyze predictor relationships
Description	Commonly employed for binary classification tasks, where the objective is to predict the probability of an event taking place by utilizing input variables. The significance level for this model is set at 0.05.
Outcome	The two clusters resulting from LCMM analysis: 0 for the first cluster & and 1 for the second cluster (binary outcome)
Predictors	Time invariant predictors
Command/Package	"glm" command from "glmnet" package in R******

Estimation of Latent Class Linear Mixed Model (LCMM): Does not use a strict 0.05 significance cutoff

Generalized Linear Mixture Model (GLMM): 0.05 significance level

Classification and Regression Trees (CART): CART models do not have a specific 0.05 significance cutoff liketraditional statistical tests

Univariate Logistic Regression: 0.05 significance level

References for Table 1:

*Magidson, J., J.K. Vermunt, and J.P. Madura, Latent class analysis. 2020: SAGE Publications Limited Thousand Oaks, CA, USA:

**Schreiber, J.B. and A.J. Pekarik, Using Latent Class Analysis versus K-means or Hierarchical Clustering to understand museum visitors. Curator: The Museum Journal, 2014. 57(1): p. 45–59

***Proust-Lima, C., V. Philipps, and B. Liquet, Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: The R Package lcmm. Journal of Statistical Software, 2017. 78(2): p. 1—56

****Douglas Bates, M., B. Bolker, and S. Walker, Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 2015. 67(1): p. 1–48

*****Boehmke, B. and B.M. Greenwell, Hands-on machine learning with R. 2019: CRC press

******Berk, R.A., Classification and regression trees (CART). Statistical learning from a regression perspective, 2008: p. 1–65

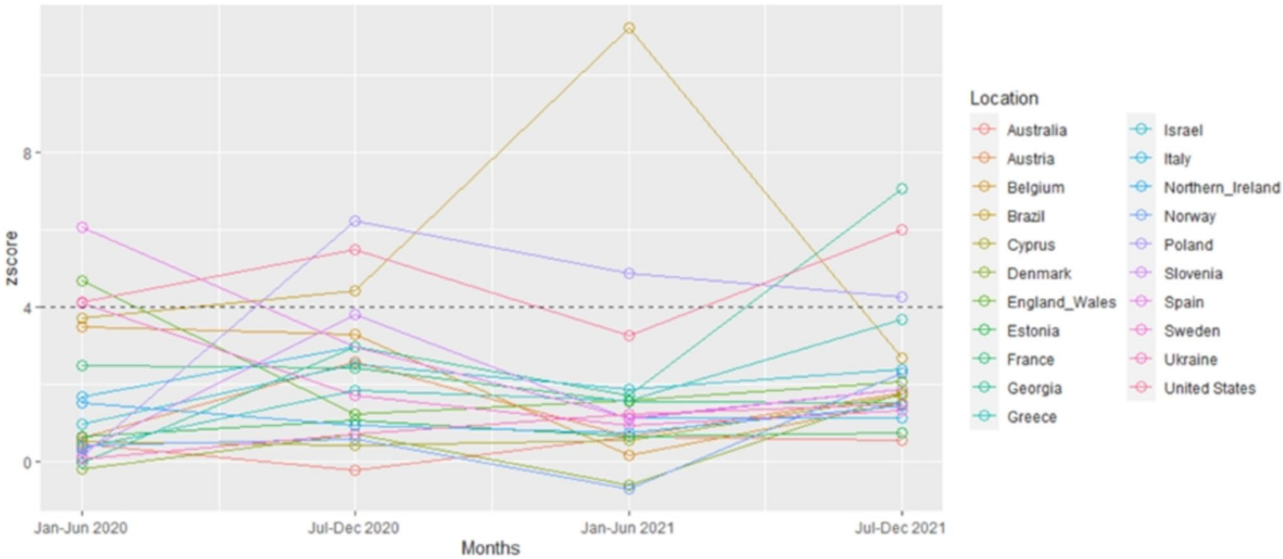

Fig. 1 Trajectory of z-score for each country/region (each point represents the means of z-score in six months). Additionally, z-score exceeding + 4 indicates substantial excess mortality. Additionally, the black horizontal line shows the z-score exceeding + 4 in

Table 2 Countries classification based on time-variant dataset

Cluster I	Mean of z-score in 2020 and 2021	Cluster II	Mean of z-score in 2020 and 2021
Australia	0.36	Brazil	5.65
Austria	1.47	Georgia	3.09
Belgium	1.99	Poland	4.19
Cyprus	0.83	United States	4.77
Denmark	0.48		
England & Wales	2.21		
Estonia	0.79		
France	1.95		
Greece	1.99		
Israel	2.01		
Italy	1.74		
Northern Ireland	1.13		
Norway	0.68		
Slovenia	1.86		
Spain	2.76		
Sweden	1.85		
Ukraine	0.94		
Mean of variable	1.47		4.43

did not yield meaningful cut-off points. To address this and ensure robust variable inclusion, a univariate regression approach was adopted to capture some relevant information within the available sample size. To run the univariate regression, countries in the first cluster were assigned a value of zero, while countries in the second cluster were assigned a value of one. Therefore, a negative value of beta would indicate that the second cluster had a lower value than the first cluster. Among all independent variables, HAQ (β: − 6.25), hypertension (β: 6.39), life expectancy (β: − 6.08), UHC (β: − 5.66), human development index (β: − 4.48), gender inequality index (β: 12.11) and government revenue (β: − 15.19) were significantly different between two clusters. The comprehensive table for univariate analysis can be found in Supplementary Table S6.

4 Discussion

Country-level responses to the COVID-19 pandemic resulted in differences in excess mortality during the pandemic. This study aimed to further examine these differences by clustering countries with similar excess mortality trajectories to identify control measures and country-level commonalities. After running the LCMM model, two clusters were extracted. Among 21 countries/regions, Brazil, the USA, Poland, and Georgia were assigned to the second cluster due to their different excess mortality trajectories during the pandemic; these countries also had the highest excess mortality rate during this period (z-score of approximately 4.4 compared to 1.5 for the first cluster). Indeed, according to Wang et al. estimation (2022), the USA and Brazil were part of the countries that experienced the highest rate of excess mortality in 2020 and 2021 [2]. Time-variant variables (and their interactions) and governmental factors (government revenue and government effectiveness) were strongly associated with excess mortality in both clusters.

Table 3 Coefficients of significant variables associated with excess mortality for countries in the first cluster

Variable	Estimate	Std. Error	Pr(>|t|)
Government effectiveness	− 0.361	0.099	< 0.001***
Unemployment	− 0.153	0.069	0.027 *
Government revenue	− 0.700	0.131	< 0.0001 ***
Obesity	− 1.196	0.453	0.039 *
Total number of nursing personnel per 1000	− 0.202	0.058	< 0.001***
Total number of medical doctors per 1000	− 0.115	0.025	< 0.0001 ***
Population over 65 years old	− 0.705	0.246	0.007 **
Stringency index (three weeks lag)	0.124	0.012	< 0.0001 ***
Proportion of population fully vaccinated with (three weeks lag)	0.143	0.019	< 0.0001 ***
COVID incidence (three weeks lag)	0.856	0.332	0.011**
Proportion of fully vaccinated (three weeks lag): COVID incidence (three weeks lag)	− 0.530	0.196	0.007 **
Stringency index (three weeks lag): proportion of fully vaccinated (three weeks lag)	− 0.115	0.039	0.003 **
Stringency index (three weeks lag): COVID incidence (three weeks lag)	− 0.780	0.399	0.051

Note: 0 \, '***' \, 0.001: The variable is highly significant at the 0.001 level

0.001 \, '**' \, 0.01: The variable is very significant at the 0.01 level

0.01 \, '*' \, 0.05: The variable is significant at the 0.05 level

0.05 \, '.' \, 0.1: The variable is marginally significant at the 0.1 level

Table 4 Coefficients of significant variables associated with excess mortality for countries in the second cluster

Variable	Estimate	Std. Error	Pr(>|t|)
Government effectiveness	− 0.342	0.162	0.035 *
Unemployment	0.172	0.074	0.02 *
Government revenue	− 2.234	0.792	0.005 **
Population over 65 years old	1.933	0.636	0.003 **
Stringency index (three weeks lag)	0.043	0.020	0.029 *
COVID incidence (three weeks lag)	8.225	0.894	< 0.0001 ***
Proportion of fully vaccinated (three weeks lag): COVID incidence (three weeks lag)	− 2.433	0.896	0.007 **
Stringency index (three weeks lag): COVID incidence (three weeks lag)	− 6.379	1.198	< 0.0001 ***

Note: 0 \, '***' \, 0.001: The variable is highly significant at the 0.001 level

0.001 \, '**' \, 0.01: The variable is very significant at the 0.01 level

0.01 \, '*' \, 0.05: The variable is significant at the 0.05 level

0.05 \, '.' \, 0.1: The variable is marginally significant at the 0.1 level

There is a possibility that the similarities in the excess mortality trajectory of Brazil, the USA, Poland, and Georgia stem from similarities in government characteristics, political contexts and COVID policies. According to policy evaluations of government responses to COVID-19, in highly polarized contexts, such as the USA and Poland intense political tensions negatively affected COVID-19 control, through reduced trust in governmental control measures [28]. The circumstances in Brazil were similarly challenging. The denial of scientific data, downplaying of the pandemic threat, and spread of fake news by the president led to uncertainty and a feeling of low risk, leading to increased exposure. The president's vaccination distribution management was problematic, and his government limited the number of vaccines received. In fact, the Parliamentary Commission of Inquiry on pandemic management uncovered purposeful delays in vaccination purchasing [29]. Georgia, on the other hand, faced political turmoil in that period. Due to the upcoming parliamentary election and a lack of evidence-based policymaking, the government removed unpopular limitations. Thus, COVID-19 infection spread swiftly in the community while the pandemic response was poorly coordinated. A near-total lockdown was implemented in late 2020 due to health system capacity saturation. Delays in immunization campaigns, poor pandemic response coordination, and a

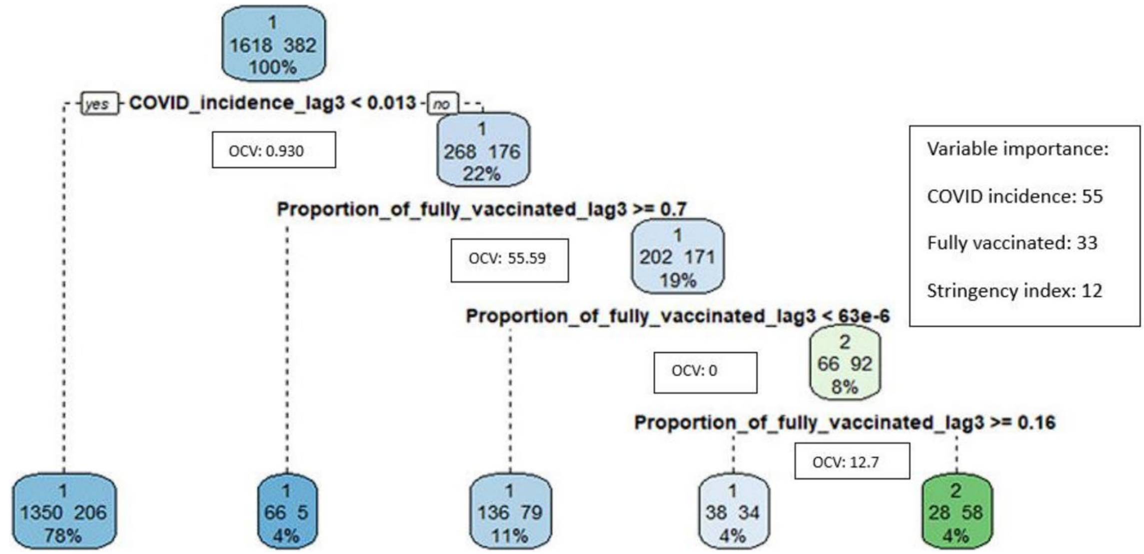

Fig. 2 Optimal tree created by CART for time-variant dataset. The blue nodes are predicted to be assigned to Cluster 1, and the green nodes are predicted to be assigned to Cluster 2. * OCV: Approximate original corresponding values of scaled data

lack of evidence-based policies led to another disastrous pandemic wave in summer 2021 [30].

The GLMM models revealed that among time-variant variables (incidence of COVID-19, fully vaccinated people, and stringency index) the incidence of COVID-19 in the population had the highest positive association with excess mortality in 2020 and 2021, while the interactions between time-variant variables mostly had a significant negative association with excess mortality. The effect estimates for the relationship between the time-variant variables and excess mortality were more pronounced in the second cluster, which countries that experienced a higher rate of excess mortality in 2020 and 2021. The lack of impact of control measure stringency and vaccination on controlling excess mortality was an unexpected result in the current study. However, this result is supported by previous studies, which show that the effectiveness of control measures may have been measure and not stringency dependent [31]. The lack of a relationship between vaccination and excess mortality was also an unexpected finding. By considering 20 countries with the highest rate of vaccination, Sub-Saharan countries and the rest of the world, a recent study found that the vaccination against COVID-19 reduced case fatality rate by an odds ratio of 0.64 [32], suggesting that the lack of a negative relationship in our study may be an artifact of the ecological nature of this investigation. Although our research found that stringency index and vaccination had a negligible positive relationship with excess mortality, the interaction between these policies and COVID-19 incidence demonstrated a stronger inverse relationship with excess mortality, possibly suggesting effectiveness of these control measures via reduced community spread. This finding highlights the significance of viewing COVID-19 policies as interdependent approaches, considering their interactions with COVID-19 incidence rather than assessing them in isolation.

Among time-invariant predictors in the GLMM models, governmental factors (government revenue and government effectiveness) had a highly negative relationship with excess mortality, highlighting the fact that countries with better government ability performed better in controlling excess mortality in this period, a finding which is aligned with other studies [33–35]. For example, by evaluating national governance in 213 countries, da Silva et al. (2023) found that Control of Corruption, Government Effectiveness, Regulatory Quality, and Rule of Law had a negative relationship with excess mortality from 2020 to 2022 [36]. Furthermore, in alignment with other studies, in the second cluster of countries in our study (with higher excess mortality), excess mortality increased with an increasing proportion of people aged 65 +. Demetriou et al. (2023) also showed that in countries such as Brazil and the USA, excess mortality in 2020 was substantially higher in the oldest population compared to other age groups [17], while a study analyzing COVID-19 fatalities and excess mortality data, adjusting for income levels across countries, revealed that individuals over 60 accounted for more than 80% of total COVID-19 deaths [37].

However, regarding the first cluster of countries (with lower excess mortality), our study observed some unexpected relationships. One was the negative relationship between obesity and excess mortality which was also observed in previous studies [38]. Another challenging result of our study is that in our Cluster 1 (with lower excess mortality), a higher rate of people aged 65 + was associated with lower excess mortality. Taken together, an increased prevalence of obesity and higher median age in a given population are possible indicators of affluence, and higher rates of comorbidities and thus heightened susceptibility among individuals. Considering the ecological nature of our study design, countries in Cluster 1 (with lower excess mortality), such as Greece, due to their higher rates of obesity and median population age may have implemented more rigorous measures of control, thus offering a plausible explanation of these findings. Affluent countries also had greater and earlier access to vaccines and were thus able to protect their more vulnerable citizens. Moreover, a relatively greater proportion of the vulnerable population, including older and obese individuals, may have been vaccinated during the initial phase of the vaccination campaign contributing to the reduced mortality.

The utilization of CART in time-variant analysis revealed that the first cluster (with lower excess mortality) exhibited a lower incidence of COVID-19 and a higher rate of full vaccination, which is consistent with anticipated results. Regarding the CART results of the time-invariant variables, the univariate logistic regression analysis reveals that countries in the first cluster (with lower excess mortality) have significantly higher HAQ, life expectancy, UHC, human development index, and government revenue, while countries in the second group (with higher excess mortality) have significantly higher hypertension and gender inequality index. In agreement with our findings, a systematic COVID-related assessment found that the HAQ index, COVID incidence, universal health coverage, aged population, and other factors contributed to nearly 70% of variations of mean level excess mortality [2]. Another study, which investigated Organization for Economic Cooperation and Development (OECD) countries, found strong negative correlations between sustainability indexes like the Sustainable Development Goals (SDG) Index, Human Development Index (HDI), and Environmental Performance Index (EPI), and COVID-19 excess mortality. This implies the potential for these indexes to serve as predictors of pandemic outcomes [29].

This study is notable for thoroughly investigating various factors contributing to excess mortality rates in the countries involved. The clustering of nations according to their excess mortality trajectories provides a unique strategy for evaluating factors associated with excess mortality within comparable trajectories across different regions. This approach differs from studies focusing on pre-established country categorizations.

However, our study is not without limitations. First, since population-level correlations may be proxies for other factors directly affecting excess deaths or death reporting, the relationships observed between excess mortality and its determinants may not represent individual experience due to possible ecological fallacy. Individual level data analysis is imperative to either strengthen some of our conclusions i.e. negative relationship between excess mortality and vaccination or shed light on other findings i.e. negative relationship between mortality and obesity. Second, when interpreting the results of the time-invariant predictors, in addition to ecological study constraints, it is important to highlight the relatively small number of countries in the second group (four countries), where lack of power may have led to uncertain results and may have compromised the generalizability of these findings. Hence, further investigation is recommended, especially with the inclusion of additional countries which experienced high excess mortality. Third, with a minimum GDP of $12,408 (Ukraine in 2020), this analysis focuses primarily on middle-high to high-income countries, thus limiting the generalizability of findings to lower income settings. Another limitation is that, as time-invariant variables mostly remain the same for both years, the number of observations for comparing two classes is limited, which could not allow researchers to assess confounding factors. Furthermore, as noted by Keep et al. (2022), it is crucial to consider that changes in population demographics across age groups from 2010 to 2019 could influence excess death estimates [39]. Despite the use of age-standardized mortality rates in this investigation, the effect of changing demographics between and within countries is an important factor to consider when interpreting the results. In addition, some countries, such as Australia, had only just opened international borders and experienced community-wide transmission very late in 2021, and so some deferred excess mortality, whether related to COVID-19 or the pandemic response, would not be captured within the time range of these analyses. Lastly, another limitation of this study is the possibility of data quality variations between countries, particularly in developing countries where data collection and reporting systems may be less robust. Similarly, registration data quality displayed temporal variability during the pandemic, even within countries, based on the strain faced by healthcare and other systems. Nevertheless, the observation of relationships across clusters of countries, rather than individual countries, adds to the study's robustness.

5 Conclusion

This study highlights the importance of COVID-19 policies, including strategies to mitigate the transmission of the virus, the strictness in implementing these measures, and the advancement of vaccination efforts, in the pandemic response. Additionally, it explores the influence of governmental elements, such as government revenue and government effectiveness, in managing the excessive deaths resulting from the COVID-19 pandemic. Furthermore, the study shows the strong influence of COVID-19 incidence on excess mortality, as the relationship between the incidence of COVID-19 and excess mortality and its interactions with other time-variant variables, were stronger in the second group, which saw higher excess mortality rates in 2020 and 2021. These findings are relevant for policymakers responsible for implementing policies during a pandemic; they highlight the importance of investing in the health protection and health resilience of older populations, the need for investment in healthcare personnel, and the importance of vaccination, in a healthy political context. Finally, this study shows that by clustering countries according to their excess mortality, researchers can acquire a deeper understanding of the impacts of national level pandemic responses.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s44197-024-00320-7.

Author Contributions Authors MRRH, CAD, CThP, SA, AQ, JQ, AA, SP, TL, AP, AC, ES and CV contributed to (1) conception and study design, or data analysis and interpretation, (2) drafting or critical revision of the article, and (3) final approval of the manuscript. Authors JACM, GD,BŁ, QH, KA, CB, CZ, WT, SNM, TH, NLM, JCPL, GA, IE, BB, JAC, LPG, OV, JTO, LM, LK, MC, JC, AJLC, TM, DA and BB contributed to (1) conception and study design, or data acquisition, or data analysis and interpretation, and (2) final approval of the manuscript.

Data Availability The data and accompanying statistical analysis codes underlying this research, in addition to what is provided in the journal and its online supplemental materials, are available upon request.

Declarations

Conflict of Interest The authors declare no competing interests.

References

1. Organization WH Global excess deaths associated with COVID-19 (modelled estimates), in Organization, World Health. 2021.
2. Wang H et al. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. The Lancet, 2022.
3. Msemburi W, et al. The WHO estimates of excess mortality associated with the COVID-19 pandemic. Nature. 2023;613(7942):130–7.
4. Davies B et al. Community factors and excess mortality in first wave of the COVID-19 pandemic in England. Nature Communications, 2021. **12**(1).
5. Kapitsinis N. The underlying factors of excess mortality in 2020: a cross-country analysis of pre-pandemic healthcare conditions and strategies to cope with Covid-19. BMC Health Serv Res. 2021;21(1):1–19.
6. da Silva RE, et al. The impact of social cohesion and risk communication on excess mortality due to COVID-19 in 213 countries: a retrospective analysis. BMC Publ Health. 2024;24(1):1598.
7. Sun MW, Troxell D, Tibshirani R. Public health factors help explain cross country heterogeneity in excess death during the COVID19 pandemic. Sci Rep. 2023;13(1):16196.
8. Watson OJ, et al. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. Lancet Infect Dis. 2022;22(9):1293–302.
9. Mendoza-Cano O, et al. Assessing the influence of COVID-19 vaccination coverage on excess mortality across 178 countries: A cross-sectional study. Vaccines. 2023;11(8):1294.
10. Flaxman S, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020;584(7820):257–61.
11. Pallari CT, et al. Magnitude and determinants of excess total, age-specific and sex-specific all-cause mortality in 24 countries worldwide during 2020 and 2021: results on the impact of the COVID-19 pandemic from the C-MOR project. BMJ Glob Health. 2024;9(4): e013018.
12. Hassan MM, et al. Assessment of epidemiological determinants of COVID-19 pandemic related to social and economic factors globally. J Risk Financial Manag. 2020;13(9):194–194.
13. Valero M, Valero-Gil JN. Determinants of the number of deaths from COVID-19: differences between low-income and high-income countries in the initial stages of the pandemic. Int J Soc Econ. 2021;48(9):1229–44.
14. Agostinis G, Parthenay K. Exploring the determinants of regional health governance modes in the Global South: a comparative analysis of Central and South America. Rev Int Stud. 2021;47(4):399–421.
15. Bouba Y, et al. The determinants of the low COVID-19 transmission and mortality rates in Africa: a cross-country analysis. Front Public Health. 2021;9:751197–751197.
16. Varkey RS, et al. Socioeconomic determinants of COVID-19 in Asian countries: an empirical analysis. J Public Aff. 2021;21(4):e2532–e2532.
17. Demetriou CA, et al. Impact of the COVID-19 pandemic on total, sex-and age-specific all-cause mortality in 20 countries worldwide during 2020: results from the C-MOR project. Int J Epidemiol. 2023;52(3):664–76.
18. Schielzeth H, et al. Robustness of linear mixed-effects models to violations of distributional assumptions. Methods Ecol Evol. 2020;11(9):1141–52.
19. Nielsen J, Krause TG, Mølbak K. Influenza-associated mortality determined from all-cause mortality, Denmark 2010/11-2016/17: The FluMOMO model. Influenza Other Respir Viruses. 2018;12(5):591–604.
20. Farrington C, et al. A statistical algorithm for the early detection of outbreaks of infectious disease. J R Stat Soc A Stat Soc. 1996;159(3):547–63.
21. Serfling RE. Methods for current statistical analysis of excess pneumonia-influenza deaths. Public Health Rep. 1963;78(6):494.
22. Hale T, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav. 2021;5(4):529–38.

23. Boehmke, B. and B.M. Greenwell, *Hands-on machine learning with R*. 2019: CRC press.
24. Yeo IK, Johnson RA. A new family of power transformations to improve normality or symmetry. Biometrika. 2000;87(4):954–9.
25. Franke, G.R., *Multicollinearity*. Wiley international encyclopedia of marketing, 2010.
26. Magidson J, Vermunt JK, Madura JP, Latent class analysis. SAGE Publications Limited Thousand Oaks. USA: CA; 2020.
27. Schreiber, J.B. and A.J. Pekarik, *Using Latent Class Analysis versus K-means or Hierarchical Clustering to understand museum visitors*. Curator: The Museum Journal, 2014. **57**(1): p. 45–59.
28. Altiparmakis A, et al. Pandemic politics: policy evaluations of government responses to COVID-19. West Eur Polit. 2021;44(5–6):1159–79.
29. Sequeira ARS, Estrela M, DeWit K COVID-19 Government policies in Portugal and Brazil: A three-year retrospective analysis. Health Policy and Technology, 2023: p. 100809–100809.
30. Nadareishvili I, et al. The Approach to the COVID-19 pandemic in Georgia—a health policy analysis. Int J Publ Health. 2022;67:1604410–1604410.
31. Herby, J., L. Jonung, and S. Hanke, *A literature review and meta-analysis of the effects of lockdowns on covid-19 mortality-II*. medRxiv, 2023: p. 2008–2023.
32. Haider N, et al. The disproportionate case–fatality ratio of COVID-19 between countries with the highest vaccination rates and the rest of the world. IJID Regions. 2023;6:159–66.
33. Gebremichael B, et al. Impact of good governance, economic growth and universal health coverage on COVID-19 infection and case fatality rates in Africa. Health Res Policy Syst. 2022;20(1):1–9.
34. Nabin MH, Chowdhury MTH, Bhattacharya S. It matters to be in good hands: the relationship between good governance and pandemic spread inferred from cross-country COVID-19 data. Humanities Soc Sci Commun. 2021;8(1):1–15.
35. Wagschal U. The influence of democracy, governance and government policies on the COVID-19 pandemic mortality. Euro Policy Anal. 2022;8(2):231–47.
36. da Silva RE, et al. National governance and excess mortality due to COVID-19 in 213 countries: a retrospective analysis and perspectives on future pandemics. Glob Health. 2023;19(1):80–80.
37. Wong MK, et al. COVID-19 mortality and progress toward vaccinating older adults—World Health Organization, worldwide, 2020–2022. Morb Mortal Wkly Rep. 2023;72(5):113–113.
38. de Siqueira JVV, et al. Impact of obesity on hospitalizations and mortality, due to COVID-19: a systematic review. Obes Res Clin Pract. 2020;14(5):398–403.
39. Kepp KP, et al. Estimates of excess mortality for the five Nordic countries during the COVID-19 pandemic 2020–2021. Int J Epidemiol. 2022;51(6):1722–32.

Authors and Affiliations

Mohammad Reza Rahmanian Haghighi[1,2] · Chryso Th Pallari[1] · Souzana Achilleos[1] · Annalisa Quattrocchi[1] · John Gabel[3] · Andreas Artemiou[4] · Maria Athanasiadou[5] · Stefania Papatheodorou[6,7] · Tianyu Liu[8] · José Antonio Cernuda Martínez[9] · Gleb Denissov[10] · Błażej Łyszczarz[11] · Qian Huang[12] · Kostas Athanasakis[13] · Catherine M. Bennett[14] · Claudia Zimmermann[15] · Wenjing Tao[16] · Serge Nganda Mekogo[17] · Terje P. Hagen[18] · Nolwenn Le Meur[19] · Jackeline Christiane Pinto Lobato[20] · Giuseppe Ambrosio[21] · Ivan Erzen[22] · Binyamin Binyaminy[23] · Julia A. Critchley[24] · Lucy P. Goldsmith[25] · Olesia Verstiuk[26] · Jideofor Thomas Ogbu[27] · Laust H. Mortensen[28] · Levan Kandelaki[29] · Marcin Czech[30] · Joseph Cutherbertson[31] · Eva Schernhammer[32] · Catharina Vernemmen[17] · Antonio José Leal Costa[33] · Tamar Maor[34] · Dimos Alekkou[35] · Bo Burström[16] · Antonis Polemitis[36] · Andreas Charalambous[37] · Christiana A. Demetriou[1]

✉ Christiana A. Demetriou
demetriou.chri@unic.ac.cy

1 Department of Primary Care and Population Health, University of Nicosia Medical School, Nicosia, Cyprus

2 Centre for Health Economic Research and Modelling Infectious Diseases, University of Antwerp, Antwerp, Belgium

3 Massachusetts Institute of Technology, Cambridge, MA, USA

4 Department of Information Technologies, University of Limassol, Limassol, Cyprus

5 Health Monitoring Unit, Government of the Republic of Cyprus Ministry of Health, Nicosia, Cyprus

6 Department of Biostatistics and Epidemiology, Rutgers School of Public Health, Piscataway, NJ, USA

7 Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, USA

8 Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA

9 Unit for Research in Emergency and Disaster, University of Oviedo, Oviedo, Spain

10 Department of Registries, National Institute for Health Development, Tallinn, Estonia

11 Department of Health Economics, Nicolaus Copernicus University in Toruń, Bydgoszcz, Poland

12 Center for Rural Health Research, East Tennessee State University, Johnson, TN, USA

13 Department of Public Health Policy, University of West Attica, Athens, Greece

14 Institute for Health Transformation, Deakin University, Victoria, Australia

15 Department of Epidemiology, Center for Public Health, Medical University of Vienna, Vienna, Austria

16 Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

[17] Epidemiology of Infectious Diseases, Sciensano, Brussels, Belgium

[18] Department of Health Management and Health Economics, University of Oslo, Oslo, Norway

[19] CNRS, Inserm, Arènes-UMR 6051, RSMS-U 1309, University of Rennes, EHESP, Rennes, France

[20] Epidemiology, Fluminense Federal University, Niterói, Brazil

[21] Medicine, University of Perugia, Perugia, Italy

[22] Public Health School, National Institute of Public Health, Ljubljana, Slovenia

[23] Israel Center for Disease Control, Ministry of Health, Ramat Gan, Israel

[24] Population Health Research Institute, City St George's, University of London, London, UK

[25] School of Health & Psychological Sciences, City St George's, University of London, London, UK

[26] Department of Medical Science, University of Nicosia, Nicosia, Cyprus

[27] Center for Research on the Epidemiology of Disasters, UCLouvain, Brussels, Belgium

[28] Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[29] Medical Statistics, National Center for Disease Control and Public Health, Tbilisi, Georgia

[30] Department of Pharmacoeconomics, Institute of Mother and Child, Warsaw, Poland

[31] Monash University Disaster Resilience Initiative, Monash University, Clayton, Australia

[32] Department of Epidemiology, Center for Public Health, Medical University of Vienna, Vienna, Austria

[33] Instituto de Estudos em Saúde Coletiva, Universidade Federal Do Rio de Janeiro, Rio de Janeiro, Brazil

[34] Israel Center of Disease Control, Ministry of Heath, Ramat Gan, Israel

[35] Department of Psychology, University of Nicosia, Nicosia, Cyprus

[36] University of Nicosia, Nicosia, Cyprus

[37] University of Nicosia Medical School, Nicosia, Cyprus