Routledge
Taylor & Francis Group

On the association between SARS-COV-2 variants and COVID-19 mortality during the second wave of the pandemic in Europe

Katarzyna Jabłońska [a], Samuel Aballéa[b], Pascal Auquier[b] and Mondher Toumi [b]

aHealth Economics and Outcomes Research Department, Creativ-Ceutical, Kraków, Poland; bFaculty of Medicine, Public Health Department, Aix-Marseille University, Marseille, France

ABSTRACT

Objective: This study aims at investigating associations between COVID-19 mortality and SARS-COV-2 variants spread during the second wave of COVID-19 pandemic in Europe.

Methods: For 38 European countries, data on numbers of COVID-19 deaths, SARS-COV-2 variants spread through time using Nextstrain classification, demographic and health characteristics were collected. Cumulative number of COVID-19 deaths and height of COVID-19 daily deaths peak during the second wave of the pandemic were considered as outcomes. Pearson correlations and multivariate generalized linear models with selection algorithms were used.

Results: The average proportion of B.1.1.7 variant was found to be a significant predictor of cumulative COVID-19 deaths within two months before the peak and between 1 January–25 February 2021, as well as of the deaths peak height considering proportions during the second wave and the pre-peak period. The average proportion of EU2 variant (S:477 N) was a significant predictor of cumulative COVID-19 deaths in the pre-peak period.

Conclusions: Our findings suggest that spread of a new variant of concern B.1.1.7 had a significant impact on mortality during the second wave of COVID-19 pandemic in Europe and that proportions of EU2 and B.1.1.7 variants were associated with increased mortality in the initial phase of that wave.

Introduction

After a year since the coronavirus infectious disease 2019 (COVID-19) outbreak has been announced as a pandemic by the World Health Organization (WHO) on 11 March 2020 [1], there is still a growing interest in monitoring the virus spread and investigating factors which can have an impact on disease mortality [2–5]. The situation seems to have worsened in countries where new variants emerged: Brazil, UK, US and South Africa [6–8]. In the UK the new 'variant of concern' (VOC) of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been identified on 20 September 2020 (B.1.1.7, or UKV)[7].

Emerging evidence suggests that VOCs may be associated with higher spread or more severe COVID-19 disease course. As reported by the Public Health England (PHE), an increased risk of hospitalization and transmissibility has been detected for VOC B.1.1.7[9]. Emerging evidence suggests that this variant may be associated with an increased risk of death [10–13]. In addition, clinical studies showed that this newly developed variant may affect the protective efficacy of naturally acquired immunity. Studies on neutralization of

convalescent sera against distinct strains showed that B.1.1.7 was harder to neutralize than the original strain, an early Wuhan-related strain of SARS-CoV-2. Neutralization titers against the B.1.1.7 variant showed a 3-fold reduction[14]. Apart from that, earlier identified variant EU2 (mutation S:447 N), observed firstly in July 2020 in western Europe, was found to be capable of increasing SARS-COV-2 infectivity [15,16]. Besides newly identified mutations, it cannot be ignored that SARS-COV-2 variants developed during the first wave, such as 20A (mutation S:D614G), developed in February 2020 and dominant during the previous wave, can also have impact on the mortality during the second wave.

Since the preliminary clinical evidence supports the hypothesis of an increased mortality associated with the B.1.1.7 VOC in the UK, a broader view on the problem is of public health interest. The real-world evidence on the actual impact of the B.1.1.7 outside the UK is still sparse. Therefore, in this study we collected country-level data on the COVID-19 mortality during the second (winter) wave of the pandemic in Europe and investigated whether the change of the

proportion of any SARS-COV-2 variant, including the B.1.1.7 and 11 other variants identified by Nextstrain up to 25 February 2021 [16,17], has an association with COVID-19 cumulative mortality or with the height of the second-wave COVID-19 mortality peak.

This study aims at detecting potential association between COVID-19 mortality and proportion of SARS-COV-2 variants through the second wave of the pandemic in Europe with the use of multivariate regression models accounting for spatial correlation. To the best of our knowledge, the problem has not been investigated so far in this broader context. This analysis enables us to fill the gap in evidence and shed light on the causes of the increased COVID-19 mortality during the second wave and through the first months of 2021 in Europe.

Materials and methods

Data collection

A total of 38 European countries were included in the analysis. The cumulative number of COVID-19 deaths during the second (winter 2020/2021) wave of COVID-19 pandemic was the primary outcome of interest. The secondary outcome of interest was the height of COVID-19 daily deaths peak during the second wave, defined as maximum daily reported number of people who died due to COVID-19 per country, considering the period from start of the second wave to 25 February 2021.

As the start of the second wave is not easily determined, we approximated this date for each country as the median date between first wave daily deaths peak height (no later than mid-June 2020) and second wave daily deaths peak height (no sooner than mid-August 2020) when a minimum number of deaths per day ±0.1 deaths per 1 mln inhabitants was observed.

Values of cumulative number of deaths and deaths peak height were divided by the number of inhabitants of a given country and reported as number of deaths per 1 million inhabitants.

The main explanatory variables of interest, assumed to have a potential association with the above outcomes, were average proportions of SARS-COV-2 sequences among identified sequences in the same period used to form Nextstrain clades. Twelve (12) clades (19A, 19B, 20A, 20A.EU2, 20B, 20 C, 20D, 20E. EU1, 20 G, 20 H/501Y.V2, 20I/501Y.V1 (B.1.1.7), 20 J/ 501Y.V3) were identified by Nextstrain from December 2019 to March 2021 in the European region [17,18].

Four time periods were considered to analyze the association between outcomes and average proportions of each clade (virus variant) through time:

- From start of the second wave up to second-wave COVID-19 daily deaths peak,
- Within two months before reaching the second-wave COVID-19 daily deaths peak,
- In the period between 1 January 2021 to 25 February 2021,
- From start of the second wave up to 25 February 2021.

Since there were between-countries differences in the timing of the pandemic course, all time periods, except the third one, are relative to the timing of the wave progression in each country.

Additional covariates considered in the analysis were: country population size, cumulative number of COVID-19 deaths during the first wave of the pandemic, cumulative number of vaccinated people up to the end of considered period, all beds capacity (number of hospital bed units), percentage of population living in metropolitan cities with more than 1 million inhabitants, percentage of population aged 65 or more, prevalence of diabetes, cancer and obesity (2017) and gross domestic product (GDP) per capita (2019). These factors were assessed as significantly impacting the risk of severe illness or mortality from COVID-19 in the literature [2,4,19,20].

Variables indicating deaths, vaccinated people and beds capacity were considered in relation to the population size of a country.

Data on COVID-19 deaths, infections and beds capacity were obtained from Institute for Health Metrics and Evaluation (IHME) on 25 February 2021 [21]. Daily number of deaths were recalculated using 7-day moving average to minimize bias related to possible reporting fluctuations.

A dataset of 3971 SARS-CoV-2 virus strains identified between December 2019 and March 2021, used to form clades of virus variants, was downloaded from the Global Initiative on Sharing All Influenza Data (GISAID) database (https://www.gisaid.org) on 12 March 2021 [22]. In particular, the dataset included information on date and country where a given strain was observed, as well as GISAID and Nextstrain clade to which a given strain was classified. The GISAID aims to rapidly share data from influenza viruses and the coronavirus causing COVID-19, including genetic sequence [22,23]. SARS-COV-2 variants' classification used in this study was developed by Nextstrain group. Nextstrain is an open-source project

providing real-time phylogenetic analysis of SARS-CoV-2 variants, grouping them into clades [16–18]. We assumed that if a strain was observed on a given date, it could be observed in a range of ±14 days from the observation date. Since the data are not reported daily, the assumption helps to avoid possible fluctuations.

Data on number of vaccinated people were taken from Our World In Data [24]. Data on population size were taken from Worldometer.com [25]. Data on the population living in metropolitan areas were downloaded from Eurostat [26]. Data on prevalence of diseases, gross domestic product (GDP) and population age were downloaded from Our World In Data and World Bank websites [27,28].

Statistical analysis

Descriptive statistics on outcomes and explanatory variables (N, mean, standard deviation, minimum, maximum) were calculated across European countries for each of considered time periods.

Pearson correlations between average virus variants proportions (raw proportions included in a range [0, 1]) and cumulative deaths and/or deaths peak height were analyzed. Then, multivariate general linear models (GLMs) with a normal distribution function and logistic link function were run using stepwise selection algorithms, to select significant variables and avoid potential bias due to relatively low sample size. A criterion of having p-value lower than 0.1 was applied for each variable to stay in and to enter a model. Variants with mean proportion across countries lower than 0.01 during a considered period were not included in the multivariate analysis. If more than one variant proportion was found significant in a model, Pearson correlations between them were verified.

Additionally, sensitivity analysis of covariates selection was performed using the genetic algorithm applied on the GLM models, with the best model selected based on the value of Akaike's Information Criterion (AIC) corrected for small sample sizes (AICC). Models with only main effects were considered.

Given a country-level analysis, Moran's I and Geary's C statistics [29,30] were produced to check the spatial autocorrelation in values of outcomes and to determine if it should be considered in final models. In that case, Gaussian form of the spatial correlation was used, controlling for countries' latitude and longitude.

For all analyses, a p-value lower than $0 \cdot 05$ was considered as statistically significant.

Analyses were performed using SAS 9.4 software. R 3.6.2 was used to apply the genetic algorithm.

Results

Descriptive statistics

Descriptive statistics of the average proportion of virus variants and outcomes for each of considered time periods are presented in Table 1. Plots presenting the change of proportion of B.1.1.7 variant through time are presented in Figure 1, and in Figure 2 for the UK.

Moran's I and Geary's C statistics for study outcomes indicated a significant spatial autocorrelation in almost all cases (Table 2), therefore all final GLMs selected with the use of selection algorithms were further accounted for the spatial autocorrelation.

Base case analysis

Pearson correlations

Significant positive correlations between the average proportion of B.1.1.7 variant of concern and cumulative number of COVID-19 deaths were observed during the time periods: from start of the second wave to the second wave peak ($0 \cdot 39$, p = $0 \cdot 017$), within two months before the second wave peak (0.32, p = 0.047) and between 1 January – 25 February 2021 (0.47, p = 0.003).

A significant negative correlation between the average proportion of 20A (S:D614G) variant and second-wave deaths peak height was observed for the period from start of the second wave up to the peak (−0.47, p = 0.002), whereas its correlation with cumulative number of deaths during that period was close to reaching the significance (−0.32, p = $0. \cdot 051$). Considering the entire second wave period, a negative correlation between the average proportion of 20A variant and second wave deaths peak height was closed to reaching the significance (−0.30, p = 0.066).

No other significant (p < 0.05) or close to significant (p < 0.1) correlations were observed for any variants within all considered time periods.

Multivariate analysis

From second wave start to the deaths peak

The average proportion of EU2 variant was found to be significant in the GLM model with stepwise selection of cumulative number of deaths during the period from second wave start to the deaths peak, with a positive estimate (1.00, p = 0.011; respectively; Table 3). Considering the same period, average proportions of B.1.1.7 and 20A variants were

Table 1. Descriptive statistics on variants proportions and outcomes.

	N	From the second wave start to the peak		During 2 months before the peak		From 1 January to 25 February 2021		From the second wave start to 25 February 2021	
		Mean (SD)	Min-Max	Mean (SD)	Min-Max	Mean (SD)	Min-Max	Mean (SD)	Min-Max
Proportion of 19A	38	0 (0 · 01)	0–0 · 04	0 (0)	0–0 · 03	0 (0 · 01)	0–0 · 03	0 (0 · 01)	0–0 · 03
Proportion of 19B	38	0 (0 · 01)	0–0 · 03	0 (0 · 01)	0–0 · 05	0 (0)	0–0 · 01	0 (0)	0–0 · 02
Proportion of 20A (S: D614G)	38	0 · 19 (0 · 19)	0–0 · 76	0 · 22 (0 · 22)	0–0 · 72	0 · 20 (0 · 18)	0–0 · 64	0 · 21 (0 · 17)	0–0 · 58
Proportion of 20A (EU2)	38	0 · 10 (0 · 16)	0–0 · 63	0 · 10 (0 · 18)	0–0 · 80	0 · 06 (0 · 11)	0–0 · 43	0 · 09 (0 · 14)	0–0 · 57
Proportion of 20B	38	0 · 28 (0 · 24)	0–1	0 · 21 (0 · 25)	0–1	0 · 12 (0 · 15)	0–0 · 59	0 · 24 (0 · 21)	0–0 · 88
Proportion of 20 C	38	0 · 01 (0 · 02)	0–0 · 07	0 · 01 (0 · 03)	0–0 · 20	0 (0 · 01)	0–0 · 07	0 · 01 (0 · 01)	0–0 · 07
Proportion of 20D	38	0 · 03 (0 · 08)	0–0 · 41	0 · 04 (0 · 12)	0–0 · 51	0 · 02 (0 · 11)	0–0 · 68	0 · 03 (0 · 09)	0–0 · 51
Proportion of 20E (EU1)	38	0 · 18 (0 · 21)	0–0 · 71	0 · 22 (0 · 25)	0–0 · 87	0 · 15 (0 · 18)	0–0 · 55	0 · 17 (0 · 19)	0–0 · 66
Proportion of 20 G	38	0 · 01 (0 · 04)	0–0 · 27	0 (0)	0–0	0 (0 · 01)	0–0 · 02	0 · 01 (0 · 04)	0–0 · 23
Proportion of 20 H/501Y.V2	38	0 (0)	0–0	0 (0)	0–0	0 · 01 (0 · 03)	0–0 · 16	0 (0 · 01)	0–0 · 04
Proportion of 20I/501Y.V1 (B.1.1.7)	38	0 · 03 (0 · 06)	0–0 · 28	0 · 07 (0 · 17)	0–0 · 64	0 · 3 (0 · 22)	0–0 · 91	0 · 09 (0 · 08)	0–0 · 39
Proportion of 20 J/501Y.V3	38	0 (0)	0–0	0 (0)	0–0	0 (0 · 01)	0–0 · 05	0 (0)	0–0 · 01
Cumulative number of deaths [per 1 mln inhabitant]	38	431 · 23 (262 · 01)	24 · 12–1022	351 · 89 (209 · 67)	24 · 12–815 · 25	327 · 04 (208 · 49)	1 · 59–860 · 31	851 · 30 (466 · 45)	52 · 59–1734
Second wave deaths peak height [per 1 mln inhabitants]	38	..	..	..	..	..	..	11 · 98 (6 · 59)	1 · 20–28 · 32

Abbreviations: SD, standard deviation; Min, minimum; Max, maximum; mln, million.
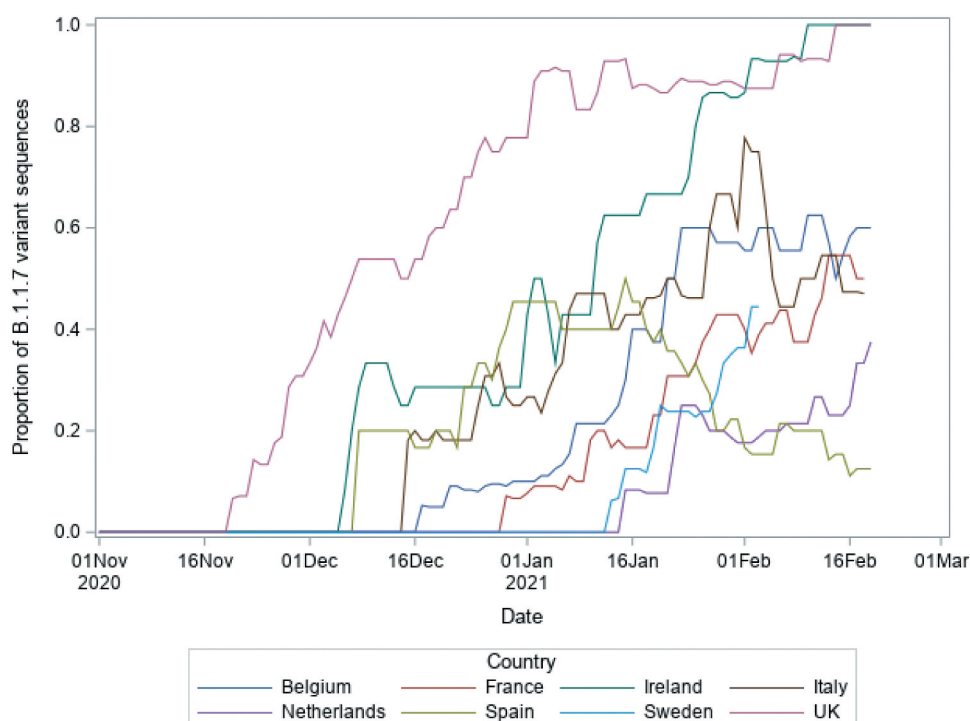
Figure 1. Proportion of B.1.1.7 variant among all variants' sequences in European countries.

found to be significantly related to the second wave deaths peak height, with a positive estimate for the former and a negative for the latter variant (3.06, p = 0.004; −1.00, p = 0.035; respectively; Table 3). It should be noted that average proportions of both variants during this period were not correlated (Pearson corr. = −0.03, p = 0.85).

Within two months before the deaths peak

For the period of two months before the peak, the average proportion of B.1.1.7 and EU2 variants were selected as significant predictors of cumulative number of deaths during that period (1.41, p < 0.001; 0.99, p = 0.001; respectively; Table 3), and the proportions were not correlated (Pearson corr. = − 0.18, p = 0.28). The B.1.1.7
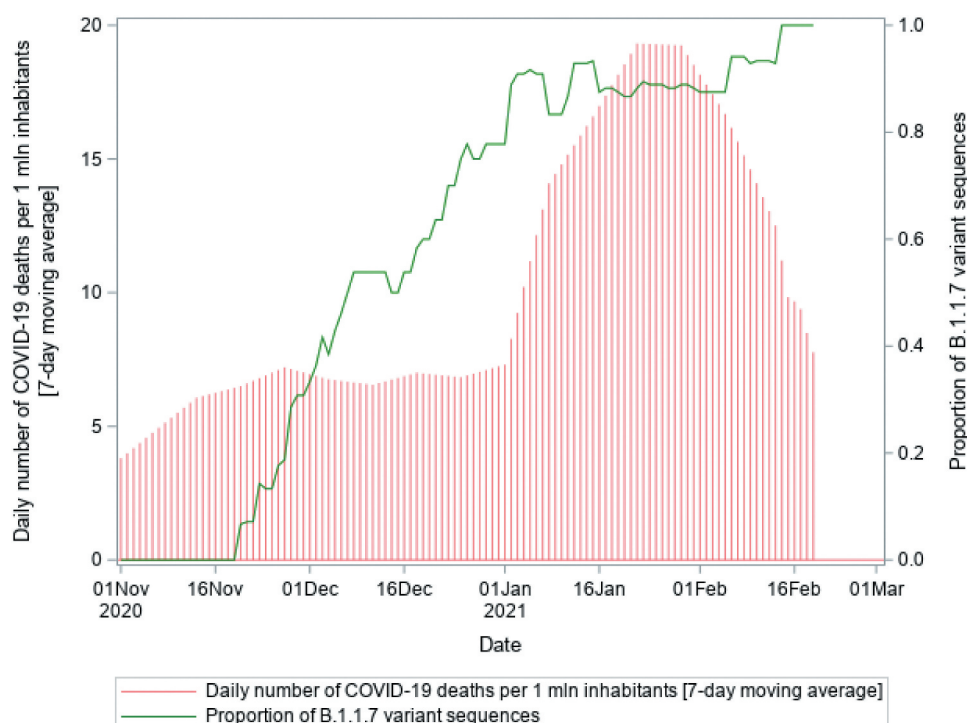
Figure 2. Proportion of B.1.1.7 variant versus daily number of COVID-19 deaths through time in the UK between 1 November 2020 and 25 February 2021.

Table 2. Spatial autocorrelation across study outcomes.

	Time period	Coefficient	Observed	Expected	SD	P value
Cumulative number of deaths [per 1 mln inhabitant]	From the second wave start to the second wave peak	Moran's I	0 · 0298	−0 · 027	0 · 0253	0.0249
		Geary's c	0 · 8201	1 · 000	0 · 0761	0.0180
	During two months before the second wave peak	Moran's I	0 · 0459	−0 · 027	0 · 0253	0.0040
		Geary's c	0 · 8949	1 · 000	0 · 0761	0.1673
	Between 1 January – 25 February 2021	Moran's I	0 · 0123	−0 · 027	0 · 0253	0.1207
		Geary's c	0 · 8178	1 · 000	0 · 0761	0.0167
	From the second wave start to 25 February 2021	Moran's I	0 · 0852	−0 · 027	0 · 0253	<0.0001
		Geary's c	0 · 8393	1 · 000	0 · 0761	0.0347
Second wave deaths peak height [per 1 mln inhabitants]		Moran's I	0.0584	−0 · 027	0 · 0253	0 · 0007
		Geary's c	0.8476	1 · 000	0 · 0761	0 · 0452

Abbreviations: SD, standard deviation; mln, million.

proportion was also found significant in the model of the deaths' peak height (1.29, p = 0.002; Table 3).

1 January – 25 February 2021
The selected GLM model of cumulative deaths for 1st of January – 25th of February 2021 includes the average proportion of B.1.1.7 variant (1.42, p < 0.001; Table 3).

From second wave start to 25 February 2021
Finally, considering the entire period of the second wave up to 25 February 2021, the average proportion of B.1.1.7 variant was selected as a significant predictor of the deaths peak height (2.37, p = 0.023; Table 3), but was not selected into the model of cumulative deaths.

Sensitivity analysis

Cumulative deaths
Considering periods from second wave start to the peak, as well as between 1 January – 25 February 2021, same models of cumulative deaths were selected by the genetic algorithm and the stepwise algorithm, suggesting their best fit based on the AICC criterion.

For the period of two months before the deaths peak, genetic algorithm selected a similar model of cumulative deaths as the stepwise algorithm but with one additional variable, percentage of people aged 65 or more. AICC of the model was only slightly better than for the stepwise model (506.70 vs 506.3).

However, proportions of B.1.1.7 and EU2 variants remained significant (1.49, p < 0.001; 0.89, p = 0.009; respectively; Table 4).

For the period from second wave start to 25 February 2021, the genetic algorithm selected a similar model of cumulative deaths as the stepwise algorithm, with GDP per capita (1 mln USD) and additionally, the percentage of people aged 65 or more (Table 4).

Deaths peak height

For all considered periods, the genetic algorithm selected the same models of the COVID-19 deaths' peak height as the stepwise algorithm.

Discussion

In this study we investigated the association between the change of SARS-COV-2 variants proportions through time and COVID-19 cumulative mortality, and the height of the second-wave COVID-19 mortality peak. The latter outcome is an indicator of mortality magnitude which can be viewed as less subject to deviations from between-country differences in reporting, not depending on the date up to which the analysis is performed and enabling to assess the overall capacity of healthcare systems [4]. Also, we decided to use data on numbers of deaths, not infections, since the former has a higher degree of consistency than the latter, being less dependent on the number of SARS-COV-2 diagnostic tests performed.

Our study provides evidence that higher proportion of the VOC B.1.1.7 across countries is associated with higher COVID-19 mortality peak and cumulative mortality during the second wave of the pandemic in Europe. An increase of 0.1 in the proportion of B.1.1.7 variant, considering the pre-peak period, was found to be associated with 35.8% increase in the height of the second wave peak. During the period from 1 January to

Table 3. Results of the GLM model using stepwise covariate selection algorithm for cumulative deaths and second wave deaths peak height, accounting for spatial correlation (N = 38).

	Estimate	Standard Error	P value
From the second wave start to the second wave peak			
Cumulative deaths during the period from the second wave start to the second wave peak			
Intercept	6 · 3101	0 · 1539	<0 · 0001
Average proportion of EU2 variant	0 · 9970	0 · 3703	0 · 0109
GDP per capita [1 mln USD]	−16 · 8039	4 · 9319	0 · 0017
Cumulative number of vaccinated people before the second wave peak [per 1 mln inhabitants]	15 · 1827	2 · 7504	<0 · 0001
Second wave deaths peak height			
Intercept	2 · 0985	0 · 5076	0 · 0002
Average proportion of 20A (S:D614G) variant	−0 · 9974	0 · 4533	0 · 0349
Average proportion of B.1.1.7 variant	3 · 0603	0 · 9726	0 · 0035
Percentage of population aged 65 or more	0 · 07303	0 · 02495	0 · 0062
Cancer prevalence	−0 · 4449	0 · 2034	0 · 0359
During two months before the second wave peak			
Cumulative deaths during two months before the second wave peak			
Intercept	5 · 9980	0 · 1405	<·0001
Average proportion of EU2 variant	0 · 9926	0 · 3220	0 · 0041
Average proportion of B.1.1.7 variant	1 · 4066	0 · 3311	0 · 0002
GDP per capita [1 mln USD]	−11 · 9994	4 · 2527	0 · 0079
Second wave deaths peak height			
Intercept	1 · 8884	0 · 5214	0 · 0009
Average proportion of B.1.1.7 variant	1 · 2865	0 · 3726	0 · 0015
Percentage of population aged 65 or more	0 · 08907	0 · 02595	0 · 0016
Cancer prevalence	−0 · 5839	0 · 2038	0 · 0071
Between 1 January – 25 February 2021			
Cumulative deaths between 1 January – 25 February 2021			
Intercept	4 · 3707	0 · 5729	<·0001
Average proportion of B.1.1.7 variant	1 · 4179	0 · 3471	0 · 0003
Percentage of population aged 65 or more	0 · 06612	0 · 02928	0 · 0304
GDP per capita [1 mln USD]	−11 · 0302	4 · 5006	0 · 0196
From the second wave start to 25 February 2021			
Cumulative deaths from the second wave start to 25 February 2021			
Intercept	7 · 0099	0 · 1302	<·0001
GDP per capita [1 mln USD]	−10 · 0637	3 · 9730	0 · 0158
Second wave deaths peak height			
Intercept	1 · 7869	0 · 5483	0 · 0025
Average proportion of B.1.1.7 variant	2 · 3707	0 · 9945	0 · 0229
Percentage of population aged 65 or more	0 · 07842	0 · 02855	0 · 0096
Cancer prevalence	−0 · 5034	0 · 2101	0 · 0222

GLM multivariate models with normal distribution and logit link function were used to explore factors associated with COVID-19 cumulative deaths and second wave deaths peak height. The number of cumulative deaths and average variants proportions were calculated during each of considered periods. Each model was run using 38 observations. Models were selected based on the use of stepwise selection algorithm.

Table 4. Sensitivity analysis results of the GLM model using genetic covariate selection algorithm for cumulative deaths, accounting for spatial correlation (N = 38).

	Estimate	Standard Error	P value
During two months before the second wave peak			
Cumulative deaths during two months before the second wave peak			
Intercept	5 · 1807	0 · 5837	<·0001
Average proportion of EU2 variant	0 · 8938	0 · 3197	0 · 0086
Average proportion of B.1.1.7 variant	1 · 4900	0 · 3368	<·0001
Percentage of population aged 65 or more	0 · 04691	0 · 03145	0 · 1453
GDP per capita [1 mln USD]	−13 · 9679	4 · 9446	0 · 0080
From the second wave start to 25 February 2021			
Cumulative deaths from the second wave start to 25 February 2021			
Intercept	6 · 2515	0 · 5010	<·0001
Percentage of population aged 65 or more	0 · 04539	0 · 02736	0 · 1060
GDP per capita [1 mln USD]	−12 · 0909	4 · 7775	0 · 0160

GLM multivariate models with normal distribution and logit link function was used to explore factors associated with COVID-19 cumulative deaths. The number of cumulative deaths and average variants proportions were calculated during each of considered periods. Each model was run using 38 observations. Models were selected based on the use of genetic selection algorithm.

25 February 2021, an increase of 0.1 in the proportion of the same VOC was related with a 15.3% increase in the cumulative number of deaths during that period.

These results support previous findings suggesting the increased risk of dying due to B.1.1.7 variant in the UK. The UK's governmental New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG) reports the results of the matched cohort analysis in which a death risk ratio for VOC-infected individuals compared to non-VOC was 1.65 (95% confidence interval (CI) = [1.21; 2.25]) [13]. Results of a case control study conducted by Challen (et al., 2021) suggest that the mortality hazard ratio associated with infection with VOC B.1.1.7 was 1.64 (95% CI = [1.32; 2. 04]) compared to infection with previously circulating variants [12]. Wallace and Ackland (2021) compared the number of deaths in the UK detected before and after new VOC detection. In early-December 2020 deaths ratio was significantly higher compared to the ratio in October-November 2021, especially in regions affected by the VOC B.1.1.7 [11].

Results of the current study also suggest that the higher proportion of EU2 variant (mutation S:447 N) was associated with increased cumulative mortality in the pre-peak phase of the second wave of COVID-19 in the European region. It is complementary to the previous findings that this variant is able to attenuate neutralizing immune responses in humans, studied by Liu et al., 2020) [15].

The proportion of 20A (mutation S:D614G) variant, firstly observed in Europe during the 1st wave of the pandemic (February 2020), was found to be negatively associated with the deaths peak height considering the pre-peak phase of the second wave. While the proportion of B.1.1.7 was positively associated with this outcome, and since no correlation between proportion of 20A and B.1.1.7 was observed in that period, then it can suggest that higher frequency of variants dominant during the first wave, such as 20A, could potentially be protective during the pre-peak phase of the second wave, due to the fact that societies could have already gained some level of immunity. Also, this finding underlines that a better understanding of how to manage the disease following the first wave may contribute to readiness for the next waves.

Limitations

One of the study limitations is that data on limited number of countries were used. Since the course of the pandemic differs sorely between continents, and so do the virus variants spread, we focused on the European region to avoid data inconsistency issues. However, the number of observations is enough to draw conclusions based on multivariate regression models. Stability of results was tested with sensitivity analyses, being highly consistent with base case findings.

The limited number of covariates were included in the multivariate analysis, while other factors could be potentially influential. The selection of covariates was based on the literature search and only factors which previously were found to have a significant association with COVID-19 mortality were included. The potential impact of seasonality on the COVID-19 mortality, previously highlighted by some authors [31,32], was not considered in this study because our analysis covered only several months across autumn and winter 2020/2021, not the full annual cycle of the disease, hence, the omitted seasonality factor should not affect the results.

Another limitation concerns the reliability and comparability of the data. The proportions of variants were calculated across strains obtained from GISAID and sampling of virus strains may not be equal across countries. However, the number of countries with less than 30 observations was only 5 (13 · 2%). We found GISAID to be a reliable source of data on SARS-COV-2 variants spread. There also exists the variability of data quality between countries, as well as between-country differences on data on daily deaths due to COVID-19 collection methods. In addition, the date of death occurrence and date of reporting can differ.

Conclusions

Our findings suggest that the development and spread of a new virus variant of concern B.1.1.7 had a significant impact on the mortality during the second wave of COVID-19 pandemic in Europe. It can contribute to explain the persistent high mortality post peak of the second wave and during the first two months of 2021.

Moreover, the frequency of both the novel VOC, as well as of earlier identified EU2 variant, could have a potential influence on excess mortality during the initial phase of the second wave, before reaching the deaths peak.

This analysis also suggests that higher frequency of variants dominant during the first wave could potentially be protective during the pre-peak phase of the second wave.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Ethical approval

Not required. This study does not require ethical approval as it was conducted on country-level data and involved information freely available in the public domain.

Authors' contribution

MT conceptualized the study; KJ, SA and PA validated the study concept. KJ and MT collected and verified the data; KJ and SA developed the methodology; KJ analysed the data; KJ, SA and MT interpreted the results; KJ wrote the first draft of the manuscript; SA, PA and MT reviewed and edited the manuscript; KJ, SA, PA and MT revised the manuscript and approved the final version.

ORCID

Katarzyna Jabłońska http://orcid.org/0000-0001-7421-443X
Mondher Toumi http://orcid.org/0000-0001-7939-7204

References

[1] World Health Organisation. Coronavirus disease 2019 (COVID-19) Situation Report – 51; 2020 March 11.
[2] Aldibasi OS, Alharbi NK, Alkelya M, et al. The association of country-level factors with outcomes of COVID-19: analysis of the pandemic after one million cases. Res Square. 2020.
[3] Garcia de Alcaniz JG, Romero-Lopez J, Martinez RP, et al. What variables can better predict the number of infections and deaths worldwide by SARS-CoV-2? Variation through time. medRxiv. 2020;2020(6):04.20122176.
[4] Jabłońska K, Aballéa S, Toumi M. Factors influencing the COVID-19 daily deaths' peak across European countries. Public Health. 2021;194:135–142.
[5] Sorci, G., Faivre, B. & Morand, S. Explaining among-country variation in COVID-19 case fatality rate. Sci Rep10, 18909 (2020). https://doi.org/10.1038/s41598-020-75848-2.
[6] Naveca F, Da Costa C, Nascimento V. SARS-CoV-2 reinfection by the new variant of concern (VOC) P.1 in Amazonas, Brazil; 2021 [cited 2021 Mar 16]. Available from: https://virological.org/t/sars-cov-2-reinfection-by-the-new-variant-of-concern-voc-p-1-in-amazonas-Brazil/596
[7] Rambaut A, Loman N, Pybus O. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. [cited 2021 Mar 15]. Available from: https://virological.org/t/preliminary-genomic-characterisationof-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563
[8] Larsen BB, Worobey M. Phylogenetic evidence that B.1.1.7 has been circulating in the USA since early- to mid-November; 2021 [cited 2021 Mar 16]. Available from: https://virological.org/t/phylogenetic-evidence-that-b-1-1-7-has-been-circulating-in-the-united-states-since-early-to-mid-november/598
[9] Public Health England. SARS-CoV-2 variants of concern and variants under investigation in England. Technical briefing 7. 11 March 2021 [cited 2021 Mar 15]. Available from: https://www.gov.uk/government/publications/investigation-of-novel-sars-cov-2-variant-variant-of-concern-20201201,https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/968581/Variants_of_Concern_VOC_Technical_Briefing_7_England.pdf
[10] Iacobucci G. Covid-19: new UK variant may be linked to increased death rate, early data indicate. BMJ. 2021;372:n230.
[11] Wallace DJ, Ackland GJ. Abrupt increase in the UK coronavirus death-case ratio in December 2020. medRxiv. 2021;2021(1):21.21250264.
[12] Challen R, Brooks-Pollock E, Read JM, et al. Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. BMJ. 2021;372:n579.
[13] New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG). Report to SAGE: increased disease severity in people infected with variant of concern (VOC) B.1.1.7 compared to people infected with non-VOC virus variants; 2021 [cited 2021 Mar 15]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/961042/S1095_NERVTAG_update_note_on_B.1.1.7_severity_20210211.pdf
[14] Supasa P, Zhou D, Dejnirattisai W, et al. Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. Cell. 2021;184(8):2201–11.e7.
[15] Liu Z, VanBlargan LA, Rothlauf PW, et al. Landscape analysis of escape variants identifies SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. bioRxiv. 2020;2020(11):06.372037.
[16] Hodcroft EB. CoVariants: SARS-CoV-2 mutations and variants of interest; 2021 [cited 2021 Mar 15]. Available from: https://covariants.org/

[17] Hadfield, Hadfield J, Megill C, et al. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics. 2018;34 (23):4121–4123.

[18] Bedford T, Neher R, Hadfield J, et al. Nextstrain - real-time tracking of pathogen evolution; 2021 [cited 2021 Mar 15]. Available from: https://nextstrain.org/

[19] Sanyaolu A, Okorie C, Marinkovic A, Patidar R, Younis K, Desai P, Hosein Z, Padda I, Mangat J, Altaf M. Comorbidity and its Impact on Patients with COVID-19. SN Compr Clin Med. 2020 Jun 25:1-8. doi: 10.1007/s42399-020-00363-4. Epub ahead of print. PMID: 32838147; PMCID: PMC7314621.

[20] Centers for Disease Control and Prevention. COVID-19: people at increased risk with certain medical conditions; 2021 [cited 2021 Mar 15]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html

[21] Institute for Health Metrics and Evaluation. COVID-19 resources - estimate downloads; 2021 [cited 2021 Feb 25]. Available from: http://www.healthdata.org/covid/data-downloads

[22] Global Initiative on Sharing All Influenza Data (GISAID). The GISAID initiative. [cited 2021 Mar 14]. Available from: https://www.gisaid.org/

[23] Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data – from vision to reality. EuroSurveillance. 2017;22(13). DOI:10.2807/1560-7917.ES.2017.22.13.30494. PMCID: PMC5388101

[24] Our World In Data. Coronavirus (COVID-19) Vaccinations. 14 March 2021 [cited 2021 Mar 14]. Available from: https://ourworldindata.org/covid-vaccinations

[25] Worldometer. Covid-19 coronavirus pandemic. [cited 2021 Feb 25]. Available from: https://www.worldometers.info/coronavirus/

[26] Eurostat. Eurostat database. [cited 2021 Jun 25]. Available from: https://ec.europa.eu/eurostat/data/database

[27] The World Bank. World Bank open data. [cited 2021 Mar 14]. Available from: https://data.worldbank.org/

[28] Our World In Data. Our world in data - research and data to make progress against the world's largest problems. [cited 2021 Mar 14]. Available from: https://ourworldindata.org

[29] Getis A, Cliff AD, Ord JK. 1973: spatial autocorrelation. London: Pion. Prog Human Geogr. 1995;19(2):245–249.

[30] Mathur M. Spatial autocorrelation analysis in plant population: an overview. J Appl Nat Sci. 2015;7:501–513.

[31] Jabłońska K, Aballéa S, Toumi M. The real-life impact of vaccination on COVID-19 mortality in Europe and Israel. Public Health. 2021;198:230–237.

[32] Liu X, Huang J, Li C, et al. The role of seasonality in the spread of COVID-19 pandemic. Environ Res. 2021;195:110874.