Comparative prevalence of COVID–19 in European countries: a time window at second wave

Abstract

Introduction: The pandemic generated by COVID–19 completely changed people’s daily lives, their relationship with family and friends, unexpectedly disrupted their working conditions and enhanced the need for an enduring resilience to face yet a second wave of the disease.

It is crucial to keep continuously updating our knowledge about COVID–19 prevalence and incidence evolutions over large connected territories, where the disease is striking in alarming proportions.

Objective: The main objective of this research is to identify and describe COVID–19 prevalence, incidence and mortality profiles in EU and EEE/EFTA countries, seven months after the start of the pandemic in Europe, and more recent tendencies, probably associated to the beginning of a second wave.

Methods: This COVID–19 study covers thirty–one European countries. Six epidemiological variables were analyzed per 100 000 inhabitants on October 25 2020, two of them evaluated over the seven previous days. A multivariate statistical exploratory analysis based on rank principal components and cluster analysis was applied.

Results: A COVID–19 prevalence typology of six country clusters was identified regarding 31 countries (EU, UK and three EEE/EFTA countries). The five epidemiological variables and number of tests revealed a wider dispersion with outlier observations. The rank transformation of data and their multivariate statistical analysis allowed us to construct a rational to better discriminate and describe these clusters, identifying specific behaviours related to the global prevalence from March until the end of October or highlight recent evolutions of COVID–19 incidence in the context of a second wave of pandemic. In fact we pinpointed country clusters where COVID–19 reached alarming levels which persist, or have even worsen, at the beginning of the second wave. Additionally, two other clusters were identified: one with countries that seems to be evolving into a situation under control, and another cluster of countries very weakly struck on the first wave, but are now facing a very complex surge, that will test their health systems capacity and timely response regarding covid and non–covid patients. Finally, the worst and more dramatic situation occurred in countries where the number of deaths per 100 000 inhabitants attained an impressive cumulative score.

Keywords: COVID–19, epidemiological, variables, rank principal components, cluster analysis

Abbreviations: EFTA, European Free Trade Association; EU, European Union; TNC, Total number of cases; TND, Total number of deaths; TS, Total number of critical serious cases; TE, Total number of tests; NWC, New cases; NWD, New deaths; RPCA, Rank principal components

Introduction

Coronaviruses, known since the 1960ies, result from the evolution of other animals which, following mutations, recombinations and adaptations, may transmitted to humans. They are therefore zoonotic strains and in this case all viruses of the RNA genome. These strains of viruses may cause infections in humans, usually respiratory infections with very diverse symptoms, varying from asymptomatic cases to very serious pseudo–flu states and even life–threatening clinical conditions.1

The disease caused by the new coronavirus SARS–Cov 2 was named COVID–19 by the World Health Organization (WHO). Its most frequent transmission pathway occurs through droplets produced in the respiratory tract that may directly reach the mouth, eyes, nose or hands of close–by people. Current evidence indicates the possibility of transmission through aerosols, hands or clothing that contacted several contaminated surfaces and subsequently touched the proximity of respiratory or ocular mucous membranes.2–4

Over the past sixty years, seven varieties of coronavirus have been identified, four of which presenting relatively mild and cold–like effects: the first, HCoV–229E in 1960, followed by HCoV–OC43 and HCoV–NL63 in 2004, and HCoV –HKU1 in 2005. The first more aggressive coronavirus appeared between 2002 and 2003 causing the acute respiratory syndrome SARS–Cov. In 2012, MERS–Cov caused the Middle East Respiratory Syndrome, and, at the end of 2019, the present outbreak evolved into an actual pandemic – SARS–Cov 2.

The most frequent symptoms associated with Covid 19 are high fever, cough and difficulty breathing, but with the increase in cases, other new symptoms, affecting other organs or systems, are
Comparative prevalence of COVID–19 in European countries: a time window at second wave

Currently, on a worldwide scale, a multitude of very different clinical trial protocols are undergoing, from which more concrete results are expected. The unprecedented scale of this effort is clear when, according to the International Federation of Manufacturers and Pharmaceutical Associations, more than 300 treatments for COVID–19 are being studied or tested on people around the world. Twenty–two pharmaceutical companies are involved in 81 clinical trials of drugs – new or old – for COVID–19, such as antivirals, anti–inflammatories, blood plasma from recovered patients and monoclonal antibodies (antibodies produced in the laboratory). Treatments for COVID–19 currently being tested are very diverse and vary depending on having a direct action on the virus or an indirect action stimulating the immune system.

Concomitantly, we are witnessing a global race in the search for an effective vaccine against COVID–19, through well–differentiated research clues. After 10 months from the onset of the disease, 213 vaccine projects are under way at various phases: 22 at clinical phase 1, 5 at phase 2 and 8 vaccines already at final phase 3. Developing a vaccine, due to its complexity, effectiveness and safety issues, usually takes much longer, for example: Polio, 20 years old, approved in 1955; Measles, 9 years, approved in 1963; Rotavirus, 22 years, approved in 2006; Malaria, 31 years, adopted in 2015; HPV, 15 years, approved in 2006; Ebola, 5 years, approved in 2019; HIV under study since 1984 and still without a vaccine; Zika virus under study since 2015 and still without a vaccine.

Regarding COVID–19, the different paths being followed to develop a future vaccine show the complexity of this research, always with the common objective of producing a satisfactory immunization within the shortest possible time:

- from the virus itself, attenuated in 4 cases or disabled in 14 cases;
- from other viruses, non–replicating in 28 cases or replicating in 26 cases;
- from genetic material, RNA in 27 cases and DNA in 17 cases;
- from proteins, subunits in 71 cases and virus–like particles in 16 cases;
- from replicative bacteria in 2 cases;
- from unknown material in 18 cases.

The COVID–19 pandemic completely changed people’s daily lives, their social and labour conditions and put a potentially overwhelming pressure on health systems, now facing a second wave of the disease.

It is thus crucial to keep continuously updating knowledge about COVID–19 prevalence, incidence and mortality evolutions over large, connected territories, namely in Europe.19,20

The main objective of this research is hence to describe COVID–19 prevalence, incidence, and mortality profiles from thirty–one countries from the European Union (EU), United Kingdom (UK), and the European Free Trade Association (EFTA), seven months after the start of the pandemic in Europe, and more recent tendencies, probably associated to the beginning of a second wave.

Methods

Data

This study includes 31 countries – EU countries, UK and three EFTA countries (Switzerland, Norway and Iceland).
Epidemiological data drawn from the European Center for Disease Prevention and Control (ECDC), referring to October, 25 2020 include the following four variables: Total number of cases, Total number of deaths, Total number of serious critical cases and Total number of tests. Additionally, two more variables were included, evaluating recent prevalence of COVID–19 covering the period October, 19–25 2020: New cases and new deaths. Indicators per 100 000 inhabitants were calculated considering 2018 population data from the World Bank, through ECDC.

**Statistical analysis**

Firstly the original data matrix was transformed into a ranks matrix, with country \(i\) \((i=1, 31)\) and variable \(j\) \((j=1,2,...,6)\) identified by \(k^{j}_{i}\), the rank of country \(i\) after the set of countries having been ranked for this variable \(j\). For each variable, countries are then ranked from 1 to 31, where 1 represents the lowest value observed and 31 the highest.

In this nonparametric approach the implicit assumptions applied to the measurements are obviously weaker, replacing the context of multinormality by just a hypothesis that data are distributed continuously. Without any ties all the variables will present the same mean value, \(\frac{n+1}{2}\), where \(n\) is the number of countries under study, and the empirical variance will be also the same for all variables, \((n^2 - 1)/12\).

The proposed transformation is thus robust to the presence of outliers, even it does not fix any asymmetric distributions.

The univariate analysis of rank data will allow us to identify, for each variable, the upper ten and lower ten countries, so the relative degree of COVID–19 prevalence ventilated by the six dimensions, namely the four first variables versus the two others referring the number of tests. Additionally, two more variables were included, including the following four variables: Total number of cases, Total number of deaths, Total number of serious critical cases and Total number of tests. Furthermore, the maximum square distance occurs when any two variables are in complete rank opposition, thus with a Spearman coefficient equal to –1. Complementary any two countries are close if their ranks are similar along the set of variables under study.

**Results**

**Univariate analysis:** A univariate statistical rank analysis of the six quantitative epidemiological variables was performed and can be found in table 1. This table presents the rank of each country \((i=1, 2, ..., 31)\) in what concerns each variable \(j\) \((j=1,..., 6)\). The top ten countries for each epidemiological variable of our study are highlighted. We may observe that until October, 25 2020, the worst prevalence of Covid–19 occurred in Spain, Belgium, Czechia, Luxemburg, France, Netherlands, Iceland, UK, Switzerland and Malta. In these countries, the total number of Covid–19 cases per 100 000 inhabitants was between 2375 in Spain and 1190 in Malta. Concerning simultaneously the serious critical cases, the most concerning situation occurred in Czechia, Belgium, Spain, Romania, France, Slovenia, Netherlands, Poland, Hungary and Portugal, varying from 6.9 in Czechia to 1.94 in Portugal. Regarding the total number of deaths by Covid–19, the top ten countries were Belgium, Spain, UK, Italy, Sweden, France, Netherlands, Ireland, Romania and Switzerland varying from 91.2 in Belgium to 23.9 in Switzerland. From Table 1 the rank scores of new cases and new deaths, give us a first look on the relative behaviours of the European countries under study at the beginning of a second Covid–19 wave. Concerning the new cases, the top ten countries were Czechia, Belgium, Switzerland, Luxemburg, Slovenia, Netherlands, Spain, France, UK and Croatia, varying from 601.4 in Czechia to 149.8 in Croatia. Finally, the new deaths between October, 19 to 25 2020, assumed a greater impact in Czechia, Hungary, Romania, Spain, Belgium, France, Poland, UK, Slovenia and Bulgaria, from 6.9 in Czechia to 1.76 in Bulgaria. We note that most of these countries had not been particularly affected by the pandemic during the first wave.

**Table 1** Rank statistical scores

|                | Total number of cases | Total number of deaths | Total serious critical cases | New cases | New death | Total tests |
|----------------|-----------------------|------------------------|-----------------------------|----------|-----------|-------------|
| Austria        | 16                    | 12                     | 20                          | 14       | 13        | 16          |
| Belgium        | 30                    | 31                     | 30                          | 30       | 27        | 26          |
| Bulgaria       | 9                     | 18                     | 21                          | 13       | 22        | 1           |
| Croatia        | 14                    | 10                     | 14                          | 22       | 21        | 3           |
| Cyprus         | 3                     | 1                      | 6                           | 7        | 1.5       | 25          |

**Citation:** Tallon JM,Gomes P,Bacelar–Nicolau L. Comparative prevalence of COVID–19 in european countries: a time window at second wave. *Biom Biostat Int J*. 2020;9(6):196–207. DOI: 10.15406/bbij.2020.09.00321
### Table Continued

| Country | Total number of cases | Total number of deaths | Total serious critical cases | New cases | New death | Total tests |
|---------|-----------------------|------------------------|----------------------------|-----------|-----------|-------------|
| Czechia | 29                    | 19                     | 31                         | 31        | 31        | 10          |
| Denmark | 12                    | 15                     | 4                          | 8         | 7         | 29          |
| Estonia | 6                     | 8                      | 1                          | 1         | 10        | 11          |
| Finland | 2                     | 9                      | 3                          | 2         | 4         | 19          |
| France  | 27                    | 26                     | 27                         | 24        | 26        | 15          |
| Germany | 8                     | 16                     | 16                         | 10        | 6         | 18          |
| Greece  | 4                     | 7                      | 9                          | 4         | 11        | 8           |
| Hungary | 10                    | 17                     | 23                         | 12        | 30        | 2           |
| Iceland | 25                    | 4                      | 10                         | 11        | 1.5       | 30          |
| Ireland | 21                    | 24                     | 8                          | 15        | 12        | 21          |
| Italy   | 15                    | 28                     | 19                         | 17        | 16        | 17          |
| Latvia  | 1                     | 3                      | 7                          | 5         | 8         | 14          |
| Lithuania | 7                    | 5                      | 11                         | 9         | 9         | 24          |
| Luxemburg | 28                  | 21                     | 15                         | 28        | 19        | 31          |
| Malta   | 22                    | 14                     | 18                         | 20        | 14        | 28          |
| Netherlands | 26                | 25                     | 25                         | 26        | 20        | 9           |
| Norway  | 5                     | 6                      | 2                          | 3         | 3         | 20          |
| Poland  | 11                    | 13                     | 24                         | 19        | 25        | 4           |
| Portugal | 20                   | 20                     | 22                         | 18        | 18        | 22          |
| Romania | 18                    | 23                     | 28                         | 16        | 29        | 7           |
| Slovakia | 13                   | 2                      | 13                         | 21        | 17        | 5           |
| Slovenia | 17                   | 11                     | 26                         | 27        | 23        | 6           |
| Spain   | 31                    | 30                     | 29                         | 25        | 28        | 23          |
| Sweden  | 19                    | 27                     | 5                          | 6         | 5         | 13          |
Bivariate and multivariate analysis: In a second step a bivariate rank analysis was performed, constructing the Spearman correlation matrix between the six epidemiological variables under study (Table 2). This table reveals a strong positive rank correlation between new cases and total cases (0.82) which means that generally the relative position of countries confirm almost the same prevalence Covid pattern registered during the first wave. However, we identified a relative worsening of the pandemic in Croatia, Poland, Slovakia, Slovenia and Switzerland and a relative flattening in Ireland, Sweden and Spain. Secondly, we underline the strong rank positive correlation between number of new deaths between October, 19 and October, 25 and the simultaneous number of serious critical patients on October, 25, which was relatively expectable, especially considering the dominant age groups that need intensive care.

In a third step a Rank Principal Components Analysis (RPCA) was applied to identify a small number of latent variables (principal components) likely to explain a significant part of the total inertia of the data. Principal components are new variables not observed that restore the main part of the initial variability of the data in descending order and exhibit the propriety of being uncorrelated variables.

The new matrix constructed by this multivariate statistical method can provide an interpretation of the rank data in a reduced dimensional space, minimizing the overall loss of variability of the initial data. In the present rank data transformation, all the epidemiological variables will have the same mean value, $\frac{n+1}{2}$, where $n$ represents the number of countries under study. The variance of such variables would be also equal under the condition of no ties in the original data. Under such condition the total inertia of data would be $p \times$ times the variance of each variables of this study. The RPCA outputs (Table 3) show the eigenvalues of the covariance matrix, by decreasing order. The first eigenvalue, $\lambda_1 = 306$, represents the variance of the first principal component, contributing 61.7% to the total inertia of data, which is the sum of empirical variances of variables under study. On the other hand, the second eigenvalue $\lambda_2 = 114$ means that second axis accounts for 23% of total inertia. The third eigenvalue $\lambda_3 = 18.5$ explains 7.7% and therefore the three first axes retained explain 92.4% of total variability. The decision of introducing a third axis is related to the fact that more than 20% of countries would not be adequately explained by taking only the two first axes. In fact, these three principal axes explain more that 80% of inertia for 84% of the countries and Austria (35%) is the only country not explained enough.

Finally, we register the very small rank correlation between the variable total number of tests and most of the remaining epidemiological variables. However, we found an exception, a negative rank correlation (−0.40) between total number of tests and new deaths. Bulgaria, Hungary, Croatia, Poland, Slovakia, and Slovenia contributed essentially to this negative correlation with a relatively smaller number of tests and relatively higher number of new cases (Table 1). This finding may be explained by the low prevalence of Covid in these countries over the past Spring.

In a third step a Rank Principal Components Analysis (RPCA) was applied to identify a small number of latent variables (principal components) likely to explain a significant part of the total inertia of the data. Principal components are new variables not observed that restore the main part of the initial variability of the data in descending order and exhibit the propriety of being uncorrelated variables.

The new matrix constructed by this multivariate statistical method can provide an interpretation of the rank data in a reduced dimensional space, minimizing the overall loss of variability of the initial data. In the present rank data transformation, all the epidemiological variables will have the same mean value, $\frac{n+1}{2}$, where $n$ represents the number of countries under study. The variance of such variables would be also equal under the condition of no ties in the original data. Under such condition the total inertia of data would be $p \times$ times the variance of each variables of this study. The RPCA outputs (Table 3) show the eigenvalues of the covariance matrix, by decreasing order. The first eigenvalue, $\lambda_1 = 306$, represents the variance of the first principal component, contributing 61.7% to the total inertia of data, which is the sum of empirical variances of variables under study. On the other hand, the second eigenvalue $\lambda_2 = 114$ means that second axis accounts for 23% of total inertia. The third eigenvalue $\lambda_3 = 18.5$ explains 7.7% and therefore the three first axes retained explain 92.4% of total variability. The decision of introducing a third axis is related to the fact that more than 20% of countries would not be adequately explained by taking only the two first axes. In fact, these three principal axes explain more that 80% of inertia for 84% of the countries and Austria (35%) is the only country not explained enough.

In a third step a Rank Principal Components Analysis (RPCA) was applied to identify a small number of latent variables (principal components) likely to explain a significant part of the total inertia of the data. Principal components are new variables not observed that restore the main part of the initial variability of the data in descending order and exhibit the propriety of being uncorrelated variables.

The new matrix constructed by this multivariate statistical method can provide an interpretation of the rank data in a reduced dimensional space, minimizing the overall loss of variability of the initial data. In the present rank data transformation, all the epidemiological variables will have the same mean value, $\frac{n+1}{2}$, where $n$ represents the number of countries under study. The variance of such variables would be also equal under the condition of no ties in the original data. Under such condition the total inertia of data would be $p \times$ times the variance of each variables of this study. The RPCA outputs (Table 3) show the eigenvalues of the covariance matrix, by decreasing order. The first eigenvalue, $\lambda_1 = 306$, represents the variance of the first principal component, contributing 61.7% to the total inertia of data, which is the sum of empirical variances of variables under study. On the other hand, the second eigenvalue $\lambda_2 = 114$ means that second axis accounts for 23% of total inertia. The third eigenvalue $\lambda_3 = 18.5$ explains 7.7% and therefore the three first axes retained explain 92.4% of total variability. The decision of introducing a third axis is related to the fact that more than 20% of countries would not be adequately explained by taking only the two first axes. In fact, these three principal axes explain more that 80% of inertia for 84% of the countries and Austria (35%) is the only country not explained enough.

Finally, we register the very small rank correlation between the variable total number of tests and most of the remaining epidemiological variables. However, we found an exception, a negative rank correlation (−0.40) between total number of tests and new deaths. Bulgaria, Hungary, Croatia, Poland, Slovakia, and Slovenia contributed essentially to this negative correlation with a relatively smaller number of tests and relatively higher number of new cases (Table 1). This finding may be explained by the low prevalence of Covid in these countries over the past Spring.

In a third step a Rank Principal Components Analysis (RPCA) was applied to identify a small number of latent variables (principal components) likely to explain a significant part of the total inertia of the data. Principal components are new variables not observed that restore the main part of the initial variability of the data in descending order and exhibit the propriety of being uncorrelated variables.

The new matrix constructed by this multivariate statistical method can provide an interpretation of the rank data in a reduced dimensional space, minimizing the overall loss of variability of the initial data. In the present rank data transformation, all the epidemiological variables will have the same mean value, $\frac{n+1}{2}$, where $n$ represents the number of countries under study. The variance of such variables would be also equal under the condition of no ties in the original data. Under such condition the total inertia of data would be $p \times$ times the variance of each variables of this study. The RPCA outputs (Table 3) show the eigenvalues of the covariance matrix, by decreasing order. The first eigenvalue, $\lambda_1 = 306$, represents the variance of the first principal component, contributing 61.7% to the total inertia of data, which is the sum of empirical variances of variables under study. On the other hand, the second eigenvalue $\lambda_2 = 114$ means that second axis accounts for 23% of total inertia. The third eigenvalue $\lambda_3 = 18.5$ explains 7.7% and therefore the three first axes retained explain 92.4% of total variability. The decision of introducing a third axis is related to the fact that more than 20% of countries would not be adequately explained by taking only the two first axes. In fact, these three principal axes explain more that 80% of inertia for 84% of the countries and Austria (35%) is the only country not explained enough.

The examination of the factor loading (Table 4) is crucial for the interpretation of the three retained factors. This table presents the linear correlation coefficients between each of the principal components and the variables under study.

### Table 2 Rank spearman correlations matrix

|                | Total cases | Total deaths | Serious critical cases | New cases | New deaths | Tests |
|----------------|-------------|--------------|------------------------|-----------|------------|-------|
| Total cases    | 1           |              |                        |           |            |       |
| Total deaths   | 0.71        | 1            |                        |           |            |       |
| Serious critical cases | 0.62  | 0.60        | 1                      |           |            |       |
| New cases      | 0.82        | 0.56         | 0.77                   | 1         |            |       |
| New deaths     | 0.55        | 0.56         | 0.85                   | 0.74      | 1          |       |
| Tests          | 0.28        | 0.10         | -0.27                  | 0.03      | -0.40      | 1     |

### Table 3 Rank Principal Components: Eigen values and Percentage of Explained Inertia

| Number | Eigenvalues | Percent |
|--------|-------------|---------|
| 1      | 306         | 61.7    |
| 2      | 113.95      | 22.98   |
| 3      | 38.3        | 7.72    |
| 4      | 19.2        | 3.87    |
| 5      | 11          | 2.29    |
| 6      | 7.54        | 1.52    |

Citation: Tallon JM, Gomes P, Bacelar–Nicolau L. Comparative prevalence of COVID–19 in european countries: a time window at second wave. Biom Biostat Int J. 2020;9(6):196–207. DOI: 10.15406/bbij.2020.09.00321
Table 4 Loading Matrix

|                           | Prin1   | Prin2   | Prin3   |
|---------------------------|---------|---------|---------|
| Rk Total Cases            | 0.84524 | 0.43857 | 0.08125 |
| Rk New Cases              | 0.90792 | 0.06935 | 0.30927 |
| Rk Total Deaths           | 0.76615 | 0.28536 | -0.56906|
| Rk New Deaths             | 0.87629 | -0.36628| -0.01847|
| Rk Serious Critical       | 0.89171 | -0.25011| 0.13291 |
| Rk Total Tests            | -0.11403| 0.95031 | 0.13867 |

From Table 4 and Figure 2 we may conclude that the first factor is positively correlated with all five epidemiological variables but not correlated with the number of tests. In fact, this first axis explains the degree of COVID–19 prevalence across countries, opposing a set of countries with lower prevalence until October, 25, or a relatively lower one when contrasting their behaviour along the first wave with the one at the beginning of second wave, to another set of countries with a higher prevalence of COVID–19 along both periods or showing a worsening situation at the beginning of second wave (Figure 1).

![Figure 1](image-url)
In addition, the second principal axis is essentially related to the behaviour of countries in what concerns the number of tests per 100,000 inhabitants, opposing countries where such number was above the rank mean value, to other countries with lower test numbers. Additionally, this axis identifies a set of countries where the total number of cases were relatively smaller from the beginning of the pandemic and knew a particular higher number of cases along the beginning of second wave, although that total number of COVID–19 cases was yet relatively smaller than the average number: Hungary, Bulgaria, Poland, Croatia, Slovakia, Slovenia and Romania (Figure 1).

We observe that Sweden and Italy are not quite well represented on the first principal plan, since the two first axes explain just 29% and 26% of Italy and Sweden inertia, respectively.

Finally, the third axis is contrasting countries which have a relative heavy number of total deaths, apparently better positioned at the second wave, with countries expressing a quite lower number of total deaths but with relative higher number of new COVID cases last October.

Considering our initial aim to characterize statistically the COVID–19 until October, 25 2020 and particularly at the beginning of the second wave, we used the factorial outputs explained by three first axes for the construction of the squared Euclidean distances matrix between any pair of countries. We applied an ascending hierarchical cluster analysis to this matrix D (31x31) (Figure 3) taking the Ward method as aggregation criteria at each interaction of the method.

The dendrogram suggested a partition of six clusters. This partition was validated using K–means non–hierarchical cluster analysis. As a matter of fact, K–means partitions are sub optional and conditioned to the initial choice of number of clusters, 6 clusters in present situation. So the final solution will be a local optima for the within–groups variability (maximizing groups homogeneity) and maximizing the between groups separability. Tables 5 to 10 characterize each cluster through the “gravity center” of the respective clusters.

The first cluster (Table 5) includes a set of 3 countries – Bulgaria, Hungary and Romania, which present a worsening of the COVID–19 situation at the beginning of the second wave with a mean value of new cases five times the observed for the third cluster, the less attained by the pandemic (203.3 versus 40.1). In the first cluster the mean value of new deaths was also the highest, more than eight times the observed in the third cluster. Additionally, the mean number of patients with serious critical COVID was the highest in cluster 1, reflecting the negative evolution of these countries since last September. Table 5 describes the rank–vectors associated to these countries, giving a clear picture of their COVID behaviours. In fact, these countries have close ranks concerning “critical serious cases” and “Total deaths” (above the global average rank number). Additionally, the mean value of new cases is around the global mean rank, but the mean value of new deaths is considerably above the average rank. Concerning the number of tests, this cluster has some of the smallest ranks, so a quite small number of tests in contrast to the other countries clusters.
Figure 3 Hierarchical clustering dendrogram.

Table 5 Cluster 1. Descriptive indicators

|                  | Total cases | Total death | Serious critical cases | New cases | New deaths | Tests    |
|------------------|-------------|-------------|------------------------|-----------|------------|----------|
| Mean             | 712.1       | 20.7        | 2.7                    | 97.2      | 2.5        | 11630.9  |
| Std Dev          | 290.8       | 10.3        | 1.2                    | 12.9      | 0.7        | 3480.6   |
| Median           | 562.3       | 15.5        | 2.1                    | 101.2     | 2.7        | 9836.5   |
| Interquartile Range | 520.4      | 18.5        | 2.2                    | 24.8      | 1.2        | 6229     |

Rank Scores

| Country   | Rank | Total cases | Total death | Serious critical cases | New cases | New deaths | Tests    |
|-----------|------|-------------|-------------|------------------------|-----------|------------|----------|
| Bulgaria  | 9    | 18          | 21          | 13                     | 22        | 1          |
| Hungary   | 10   | 17          | 23          | 12                     | 30        | 2          |
| Romania   | 18   | 23          | 28          | 16                     | 29        | 7          |

Citation: Tallon JM, Gomes P, Bacelar–Nicolau L. Comparative prevalence of COVID–19 in european countries: a time window at second wave. Biom Biomath Int J. 2020;9(6):196–207. DOI: 10.15406/bbij.2020.09.00321
The second cluster (Table 6) includes Croatia, Poland, Slovakia, Slovenia and Czechia and their common behaviour is related to the relatively high number of “new cases” and “new deaths” and, again, a smaller number of tests. It is interesting that in spite of these common points, Czechia exhibits a quite special behaviour with a very high rank score for four variables under study. This particular behaviour is quite clear on first principal plan and on the dendrogram (Figures 1 and 2).

Table 6 Cluster 2, Descriptive indicators

|                | Total cases | Total death | Serious critical cases | New cases | New deaths | Tests  |
|----------------|-------------|-------------|------------------------|-----------|------------|--------|
| Mean           | 1045.0      | 10.5        | 2.9                    | 267.0     | 2.8        | 13470.0|
| Std Dev        | 669.9       | 5.7         | 2.4                    | 198.1     | 2.4        | 3297.3 |
| Median         | 774.1       | 10.4        | 2.3                    | 149.8     | 1.8        | 12170.1|
| Interquartile Range | 927.7 | 8.4 | 3.9 | 305.7 | 3.1 | 5961.6 |

Rank Scores

| Country       | Rank Score | Rank Score | Rank Score | Rank Score |
|---------------|------------|------------|------------|------------|
| Croatia       | 14         | 10         | 14         | 22         | 21         | 3 |
| Poland        | 11         | 13         | 24         | 19         | 25         | 4 |
| Slovakia      | 13         | 2          | 13         | 21         | 17         | 5 |
| Slovenia      | 17         | 11         | 26         | 27         | 23         | 6 |
| Czechia       | 29         | 19         | 31         | 31         | 31         | 10 |

The third cluster (Table 7) includes a set of nine countries – Cyprus, Lithuania, Denmark, Finland, Norway, Germany, Estonia, Greece and Latvia. These countries have the smallest mean value for all five epidemiological variables and the second best in what concerns the mean number of tests per 100 000 inhabitants. Within this cluster the three last countries didn’t seem to give special priority to COVID–19 tests, ranked under the global average number, which could be explained by the low prevalence during the first wave.

Table 7 Cluster 3, Descriptive indicators

|                | Total cases | Total death | Serious critical cases | New cases | New deaths | Tests  |
|----------------|-------------|-------------|------------------------|-----------|------------|--------|
| Mean           | 353.8       | 6.2         | 0.5                    | 40.1      | 0.3        | 31967.3|
| Std Dev        | 140.4       | 3.6         | 0.4                    | 18.2      | 0.3        | 20248.8|
| Median         | 822.6       | 5.4         | 0.4                    | 40.3      | 0.4        | 24570.8|
| Interquartile Range | 149.3 | 5.5 | 0.8 | 37.7 | 0.5 | 15284.5 |

Rank Scores

| Country       | Rank Score | Rank Score | Rank Score | Rank Score |
|---------------|------------|------------|------------|------------|
| Cyprus        | 3          | 1          | 6          | 7          | 1.5        | 25 |
| Lithuania     | 7          | 5          | 11         | 9          | 9          | 24 |
| Denmark       | 12         | 15         | 4          | 8          | 7          | 29 |
| Finland       | 2          | 9          | 3          | 2          | 4          | 19 |
| Norway        | 5          | 6          | 2          | 3          | 3          | 20 |
| Germany       | 8          | 16         | 16         | 10         | 6          | 18 |
| Estonia       | 6          | 8          | 1          | 1          | 10         | 11 |
| Greece        | 4          | 7          | 9          | 4          | 11         | 8 |
| Latvia        | 1          | 3          | 7          | 5          | 8          | 14 |

Citation: Tallon JM, Gomes P, Bacelar–Nicolau L. Comparative prevalence of COVID–19 in european countries: a time window at second wave. Biom Biostat Int J. 2020;9(6):196–207. DOI: 10.15406/bbij.2020.09.00321
The fourth cluster (Table 8) includes seven countries – France, Netherlands, Spain, Belgium, Portugal, Switzerland and Austria, and registers the highest mean value of COVID–19 cases, serious critical cases and new cases. In these countries the pressure over National Health Systems was particularly strong by the end of October.

### Table 8 Cluster 4. Descriptive indicators

|            | Total cases | Total death | Serious critical cases | New cases | New deaths | Tests   |
|------------|-------------|-------------|------------------------|-----------|------------|---------|
| Mean       | 1572.6      | 45.1        | 3.0                    | 252.0     | 1.7        | 26176.2 |
| Std Dev    | 596.6       | 29.5        | 1.4                    | 131.6     | 0.7        | 7767.9  |
| Median     | 1588.6      | 40.6        | 2.9                    | 228.4     | 1.6        | 22681.0 |
| Interquartile Range | 1223.9 | 52.0 | 2.6 | 214.0 | 1.4 | 13252.1 |

**Rank Scores**

- **France**: 27 26 27 24 26 15
- **Netherlands**: 26 25 25 26 20 9
- **Spain**: 31 30 29 25 28 23
- **Belgium**: 30 31 30 30 27 26
- **Portugal**: 20 20 22 18 18 22
- **Switzerland**: 23 22 17 29 15 12
- **Austria**: 16 12 20 14 13 16

Analyzing the rank scores we immediately verify that some of them attained relatively alarming values. However the three last countries of this cluster (Table 8) exhibited a relatively smaller prevalence of COVID–19, where some rank scores are closer to mean values.

The fifth cluster (Table 9) includes just four countries – Ireland, UK, Italy and Sweden, being the cluster with the highest number of death per 100 000 inhabitants.

### Table 9 Cluster 5. Descriptive indicators

|            | Total cases | Total death | Serious critical cases | New cases | New deaths | Tests   |
|------------|-------------|-------------|------------------------|-----------|------------|---------|
| Mean       | 1058.1      | 55.9        | 1                      | 108       | 55.9       | 30174.1 |
| Std Dev    | 179.6       | 12.3        | 0.6                    | 45.9      | 12.3       | 11265.3 |
| Median     | 1104.1      | 60          | 0.9                    | 111.9     | 60         | 27187.4 |
| Interquartile Range | 320.4 | 21.4 | 1.1 | 86 | 21.4 | 20739.2 |

**Rank Scores**

- **Ireland**: 21 24 8 15 12 21
- **UK**: 24 29 12 23 24 27
- **Italy**: 15 28 19 17 16 17
- **Sweden**: 19 27 5 6 5 13
Comparative prevalence of COVID–19 in European countries: a time window at second wave

However, Sweden presents a quite particular behaviour within this cluster, concerning the low rank score of “new cases” and “new deaths”, so a quite smaller number of COVID–19 cases and associated deaths along the period October, 19 until October, 25 2020. Finally, the sixth cluster (Table 10) includes just three countries – Malta, Iceland and Luxemburg. These countries have the second highest mean value of COVID–19 cases and the highest number of tests. Considering the small population of these countries the epidemiological rank scores are very sensitive to any local change.

**Table 10 Cluster 6, Descriptive indicators**

|                  | Total cases | Total death | Serious critical cases | New cases | New deaths | Tests |
|------------------|-------------|-------------|------------------------|-----------|------------|-------|
| **Mean**         | 1497.4      | 12.2        | 1.2                    | 176.2     | 0.8        | 108289.3 |
| **StDev**        | 473.4       | 9.7         | 0.3                    | 122.5     | 0.7        | 43803.1  |
| **Median**       | 1259.9      | 19.2        | 0.5                    | 238       | 1.4        | 85808    |
| **Interquartile Range** | 852.8       | 19.2        | 0.5                    | 238       | 1.4        | 85808    |

**Discussion**

The objective of the cross-sectional study was to describe the COVID–19 prevalence at the end of October 2020 and some recent tendencies in Europe, associated to the beginning of a second wave.

The epidemiological variables under study were: Total number of COVID–19 cases, simultaneous critical serious cases, total number of deaths, total number of tests, referred at October, 25 2020 and two other indicators, total new cases and total new deaths, accumulated values from October, 19 until October, 25 2020.

All five epidemiological variables were positively correlated (Spearman Correlation) and not correlated with variable tests, except “new deaths” that registered a negative correlation with variable tests. Bulgaria, Hungary, Croatia, Poland, Slovakia and Slovenia contribute to this negative correlation, which maybe partially explained by the low COVID–19 prevalence registered in these countries last spring. In order to obtain a typology of countries that explained their COVID–19 prevalence, six clusters were obtained from the non–hierarchical cluster analysis and subsequent partition into classes. The first cluster and second cluster include countries where the epidemic has reached a severe clinical situation, along the October week under observation, explained by the high mean deaths, in clear contrast with their favourable performance along the first wave. Additionally, countries from the second cluster registered a relatively lower number of tests per 100 000 inhabitants.

The third cluster includes nine countries, most of them of Northern Europe, which had the smallest mean value for all epidemiological variables, namely along the second wave.

The fourth cluster represents seven countries with the highest mean value of COVID–19 cases, serious critical cases and new cases. In these countries the pressure over the National Health Systems was very high by the end of October, and particularly alarming for some of these countries (Figure 1).

The fifth cluster with just four countries had the highest mean value of total deaths. However, Sweden presented a particular behaviour considering the relatively lower number of new cases and new deaths, apparently more confident at the beginning of the second wave.

Finally, the sixth cluster includes just three small countries which registered the second higher mean value of COVID–19 cases and also a particularly high number of tests per 100 000 inhabitants.

Our overall data analysis confirms that a great part of European countries is suffering an increasing number of COVID cases, which may justify the implementation of new control measures by national, regional and local governments. However, the diversity of country profiles may indicate these measures should be taken also in a differentiated way and that some country clusters may learn from other country clusters’ experiences.

**Acknowledgments**

None.

**Conflicts of interest**

None.

**References**

1. Aronson JK. Coronaviruses – a general introduction. CEBM – Centre for Evidence–Based Medicine. 1ª, 2020.
2. Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med. 2020;382:1708–1720.
3. Santarpia JL, Rivera DN, Herrera V, et al. Transmission potential of Sars–CoV–2 in viral shedding observed at the university of nebraska medical center. medRxiv. 2020;1–12.
4. Doremalen N van, Bushmaker T, Morris DH, et al. Aerosol and surface stability of SARS–CoV–2 as compared with SARS–CoV–1. N Engl J Med. 2020;382(16):1564–1567.

Citation: Tallon JM, Gomes P, Bacelar–Nicolau L. Comparative prevalence of COVID–19 in European countries: a time window at second wave. Biom Biostat Int J. 2020;9(6):196–207. DOI: 10.15406/bbij.2020.09.00321
5. Mole B. Don’t Panic: The comprehensive ars technica guide to the coronavirus. 2020.
6. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult in patients with COVID–19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054–1062.
7. Levi M, Scully M. How I treat disseminated intravascular coagulation. Blood. 2018;131(8):845–854.
8. Zhang H, Zhou P, Wei Y, et al. Histopathologic changes and SARS–CoV–2 immunostaining in the lung of a patient with COVID–19. Ann Intern Med. 2020;M20–0533.
9. Tang N, Li D, Wang X, et al. Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. J Thromb Haemost. 2020;18:844–847.
10. Giannisi D, Ziogas IA, Gianni P. Coagulation disorders in coronavirus infected patients: COVID–19, SARS–CoV–1, MERS–CoV and lessons from the past. J Clin Virol. 2020;127:104362.
11. Siddiqi HK, Mehra MR. COVID–19 Illness in native and immunosuppressed states: a clinical–therapeutic staging proposal. J Heart Lung Transplant. 2020.
12. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult in patients with COVID–19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054–1062.
13. Wu Z Me, Googan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID–19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. JAMA. 2020;323(13):1239–1242.
14. Mehta P, McAuley DF, Brown M, et al. COVID–19: consider cytokine storm syndromes and immunosuppression. Lancet. 2020;395:1033–1034.
15. Li G, Clerq ED. Therapeutic options for the 2019 novel coronavirus (2019–nCoV). Nat Rev Drug Discov. 2020;19(3):149–150.
16. Cascella M, Rajnik M, Cuomo A, et al. Features, evaluation and treatment coronavirus (COVID–19). In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing. 2020.
17. Tang N, Bai H, Chen X, et al. Anticoagulant treatment is associated with decreased mortality in severe coronavirus disease 2019 patients with coagulopathy. J Thromb Haemost. 2020.
18. Watson RA, Johnson DM, Dharia RN, et al. Anti–coagulant and anti–platelet therapy in the COVID–19 Patient: A best practices quality initiative across a large health system. Hosp Pracr (1995). 2020;48(4):169–179.
19. Tallon JM, Gomes P, Bacelar–Nicolau L. Profiling European countries on COVID–19 prevalence and association with non–pharmaceutical interventions. Biom Biostat Int J. 2020;9(4):118‒130.
20. Nogueira PJ, de Araújo Nobre M, Costa A, et al. The Role of Health Preconditions on COVID-19 Deaths in Portugal: Evidence from Surveillance Data of the First 20293 Infection Cases. J Clin Med. 9(8):2368.
21. European Centre for Disease Prevention and Control. COVID–19 Pandemic, 2020.
22. Lebart L, Piron M, Morineau A. Statistique Exploratoire Multimensionnelle 4ème edition. Dunod 2006.

Citation: Tallon JM, Gomes P, Bacelar–Nicolau L. Comparative prevalence of COVID–19 in european countries: a time window at second wave. Biom Biostat Int J. 2020;9(6):196–207. DOI: 10.15406/bbij.2020.09.00321