Clinical Paper

Profiling European countries on COVID–19 prevalence and association with non–pharmaceutical interventions

Abstract

Introduction

It is essential to understand, on a large geographical scale, the dimension of the COVID–19 pandemic by identifying the most affected countries, knowing that all the world is suffering an unusual disruption regarding several health impacts, but also heavy economic, financial and social effects. A key role is reserved to Data Science to understand the present and to deepen a prospective analysis at COVID–19 day after.

Objective

The main objective of the present study is to describe the COVID–19 prevalence in EU and five other OECD countries using five epidemiological variables. Secondly their association with non–pharmaceutical measures taken in some countries to control and attenuate the evolution of the epidemic was analyzed.

Methods

The COVID–19 study covers twenty–six EU countries and additionally Switzerland, Norway, Turkey, Israel and United Kingdom. Five epidemiologic variables were analyzed by 100.000 inhabitants at the beginning of May 2020: total number of cases, total number of deaths, total number of active cases, total number of critical or serious cases and total number of tests. Also, eight non–pharmaceutical measures were selected for association purposes. A multivariate statistical exploratory approach with principal components, hierarchical and non–hierarchical (k–means) cluster analyses was applied.

Results

A COVID–19 prevalence typology of four country clusters was identified regarding EU countries and five OECD countries on early May. In the two clusters, with a total of ten countries where the pandemic seemed to evolve more seriously, different patterns regarding the number of tests are observed. Two other clusters, with 12 and 9 countries, show an intermediate or low prevalence but differences in testing patterns.

For EU countries of both clusters more affected, COVID–19 containment strategies were studied considering three modalities of implementation timing for eight non–pharmaceutical measures. The three different behaviors mirrored the clusters findings. Countries previously classified into cluster 1 appear together again, as do countries belonging to cluster 2. In spite of a common behavior for some measures, generally countries of cluster 2 implemented other interventions later in time. Sweden is a "special case", taking just a few of these measures, most of them later than other countries.

Keywords: COVID–19, epidemiological variables, non–pharmaceutical measures

José M. Tallon,[1,2] Paulo Gomes,[3] Leonor Bacelar–Nicolau[3,4]

[1]Sports Sciences Department, Exercise and Health, Universidade de Trás–os–Montes e Alto Douro, Portugal (UTAD)
[2]Medical Clinic Doctor Tallon
[3]NOVA IMS Information Management School, Portugal
[4]Faculdade de Medicina, Institute of Preventive Medicine and Public Health & ISAMB–Institute of Environmental Health, Universidade de Lisboa, Portugal

Correspondence: Jose Maria Tallon, Sports Sciences Department, Exercise and Health, Universidade de Trás–os–Montes e Alto Douro, Vila Real, Portugal (UTAD), Tel 00351968011657,
Email josemariatallon@clinicadrtallon.com

Received: June 22, 2020 | **Published:** July 14, 2020

Abbreviations: TNC, total number of cases; TND, total number of deaths, TA, total number of active cases; TS, total number of critical serious cases; TE, total number of tests

Introduction

The corona virus, known since the 1960ies are zootonic strains of RMA genome viruses, all resulting from an evolution of other animals and that, following some mutations, recombinations and adaptations, may be transmitted to humans. This virus strain may cause infections in humans, usually respiratory, but with a symptomatology that may vary form asymptomatic cases to pseudo flu states, more or less serious. On limit situations, the clinical framework may evolve to an acute respiratory insufficiency, pneumonia and even be life threatening.

COVID–19 was the name given by WHO to the disease caused by the new SARS–Cov2 coronavirus. The transmission more usually occurs through droplets produced in the respiratory tract, that may reach the mouth, eyes, nose or hands of bystanders. Current evidence does not exclude transmission through hands and clothes having been in contact with droplets on various surfaces, that previously touched the face in proximity of respiratory or ocular mucous membranes.[1]

Over the last 60 years, seven coronaviruses varieties were identified, four of them with relatively lighter effects similar to common colds: the first in 1960, the HCov–229E, followed by

HCov–OC43 and HCov–NL63 in 2004 and HCov–HKU1 on the following year. The first more recent aggressive coronavirus, SARS–Cov, appears between 2002 and 2003, causing an acute respiratory syndrome. In 2012 occurs the second, MERS–Cov, responsible for the Middle East Respiratory Syndrome, and, at the end of 2019, the present SARS–Cov 2 outbreak became a real pandemic.

The most common symptoms associated with COVID–19 are high fever, cough and difficulty breathing. Additionally, new manifestations that affect different organs or systems are increasingly being described, such as myalgia, headache, odynophagia, conjunctivitis, nasal congestion, anosmia and dysgeusia, heart problems, kidney failure and gastrointestinal problems. However, about 80% of diagnosed patients are asymptomatic or have mild symptoms, 15% are more severely affected and will need oxygen and about 5% have a critical clinical condition, requiring hospitalization in intensive care.

The severity of the disease is enhanced by comorbidities resulting from old age or chronic diseases such as cardiovascular or pulmonary diseases, diabetes, changes in the immune response and overweight or obesity, among others,[2] as well as by asymptomatic patients, who seem to transmit the disease at a lesser level, or presintomatics, who stand in the period of time when the virus is totally asymptomatic, but who are transmitters of the disease. The average incubation period in quarantined patients was estimated at 5.1 days. The WHO gives us a window of 2 to 14 days, but several studies state that in 1% of cases, symptoms will only develop after 14 days of active monitoring or quarantine.[3]

Currently, there is a broad scientific consensus that COVID–19's severe clinical conditions are associated with intravascular coagulation problems that can precipitate the patient's death.[4]

In the context of the evolutionary development of vertebrates, about 430 million years ago, a mechanism capable of stopping haemorrhages through the acquired blood clotting ability appeared. However, at older ages or in the context of the human organism's response to certain pathologies, there is sometimes an exaggerated coagulant reaction that explains that thrombosis is a more frequent cause of death than haemorrhage itself.[5]

COVID–19 is a viral disease associated with a systemic inflammatory response, acute interstitial pneumonia and mechanisms of activation of coagulation, both at the pulmonary level and in the other organs. In fact, the targets of this virus are the glandular epithelia of the respiratory and digestive tracts, although the virus also has tropism for vascular endotheliums.[6–8]

We can identify three phases of COVID–19:

1– Infection phase, the step of the viral response;

2– Pulmonary phase, a step of trying to balance the viral response and the inflammatory response of the organism where the thrombotic process begins;

3– Phase of hyper–inflammation, where the exaggerated response of the organism worsens the clinical framework.

During these phases, the disease tends to worsen according to certain easily measurable analytical parameters. Among them, the D–Dimer indicates the existence of abnormal levels of fibrin, that is, the formation of blood clots in the arteries or veins, which are likely to generate a severe clinical framework or even the patient's death. These

facts had already been reported on January 31 by pathologists from hospitals in Wuhan and later confirmed in Italian, Dutch and North American patients. Therefore, the detection of thrombus formation is one of the main markers that reveals a poor COVID–19 prognosis.[9]

However, in reality, Science does not yet fully understand the mechanisms that explain the lethality of this virus and how exactly it affects different organs and systems. Is pneumonia prevalent? Or blood clots? Or is it essentially a massive reaction from the immune system? In fact, the virus multiplies inside the cells and in the most severe cases the human body responds with a massive discharge of inflammatory cytokines released through the overactivation of white blood cells, causing high fever, blood clots, hypotension, lack of oxygen, acidification blood and pulmonary edema. In these situations, the clinical picture may even progress to the point where the white blood cells attack indiscriminately, reaching the healthy tissue itself and causing pulmonary, cardiac, hepatic, intestinal, renal and even genital insufficiency. This in turn can substantiate the Multiple Organ Dysfunction Syndrome (MODS), which, if not controlled in a timely manner, will facilitate the deposit of proteins and dead cells in the lung tissue, preventing oxygenation and hindering inexorably life.[10,11]

Currently, there is no proven effective treatment for COVID–19. In the early or mild stages of the disease, symptomatic treatment with anti–inflammatories and antipyretics has been used. In the most advanced stages, depending on the patient's general clinical condition, antimalarial, antiviral treatments of various types are used, as well as immunological modulators, whether corticosteroids, monoclonal antibodies or blood plasma transfusions of recovered patients. In patients hospitalized with COVID–19 in intensive care, antithrombotic strategies are part of the action protocol, with low molecular weight heparins with prophylactic or intermediate dosages, considered indicated when the values of the D–Dimer parameter are higher than normal values. Therapeutic anticoagulation is reserved for cases in which a local or systemic thrombotic pathology is diagnosed.[12–14]

Presently, very wide clinical trial protocols are being done worldwide, from which more concrete results are expected.

Concomitantly, we are witnessing a global race in the search for an effective vaccine for COVID–19, through well–differentiated research clues. The complexity of this process and the high level of safety required do not foresee a short–term availability of this active substance.

The COVID–19 outbreak was officially declared a pandemic by WHO on March 11, 2020, hitting Europe more strongly after February/March 2020. Over this period many European countries implemented non–pharmaceutical interventions (NPIs) to try and contain the effects of the epidemic. The first cycle of confinement extends roughly until the beginning of May (May 2/4), after which less strict measures begun to be taken gradually.[15]

The main aim of this paper is to study the COVID–19 prevalence in EU countries and European OECD countries, over the period February 25–May 4 2020, corresponding to the first cycle of confinement identifying higher prevalence clusters and associating them with non–pharmaceutical measures typologies.

Future studies will further analyse the following cycles, as less restrictive measures will be undertaken regarding the mobility of populations, economic activity and more advanced deconfinement phases.

Given the framework of a high contamination outbreak with great uncertainty regarding the length of virus survival, the viral load and the mutation rate, as well as the most reliable data available, five variables were selected to characterize the epidemiological impact of the virus on populations and health systems: Total number of cases (TNC); Total number of deaths (TND); Total number of active cases (TA); Total number of critical serious cases (TS) and Total number of tests (TE). To assure the comparability between countries, these indicators were analysis per 100.000 inhabitants (given the 2018 populations country censuses).

Possible associations were analysed between these epidemiological indicators and a set of eight non–pharmaceutical measures taken by European governments to slow the pandemic evolution: Stay–at–home (Orders–enforced, Recommendations for the general population, Recommendations for risk groups); Closure of educational institutions (Secondary schools/higher education, Primary schools, Daycare), Closure of public places, Mass gathering cancellations.[15]

Methods

Data

This study includes 32 countries – EU countries and five OECD surrounding countries (Israel, Norway, United Kingdom, Switzerland and Turkey).

Epidemiological data drawn from Worldometer referring to May 4th,[16] and the European Centre for Disease Prevention and Control, referring to May 2nd include the following five variables: Total number of cases, Total number of deaths (both until 02 May 2020), Total number of active cases, Total number of critical serious cases and Total number of tests (all until 04 May 2020). Indicators per 100.000 inhabitants were calculated, considering 2018 population data from the World Bank, through ECDC.[17]

Information on non–pharmaceutical interventions to prevent a large upsurge and their implementation dates were taken from ECDC's Rapid Risk Assessment and take into account the following eight measures: Stay–at–home (Orders–enforced, Recommendations for the general population, Recommendations for risk groups); Closure of educational institutions (Secondary schools/higher education, Primary schools, Daycare), Closure of public places, Mass gathering cancellations.[15] This information was considered for the following nine countries, with higher COVID–19 prevalence: Belgium, France, Ireland, Italy, Netherlands, Portugal, Spain, Sweden and United–Kingdom. Each non–pharmaceutical intervention was recoded into 3 categories: measure not taken; measure taken before March 17, 2020; measure taken on March 17, 2020 or after.

Statistical analysis

Firstly, a univariate statistical analysis of the epidemiologic variables was performed using location and dispersion statistics, quantiles and a Box–plot analysis to visualize the empirical distribution of each variable and identify lowest and highest behaviors, namely possible outliers.

Secondly, a bivariate statistical analysis interpreting the Pearson correlation matrix was undertaken to assess the statistical linear association between any pair of the epidemiologic variables.

Then, a Principal Component Analysis (PCA) was executed to identify a small number of latent variables (principal components), linear combinations of the initial centered variables, likely to explain a significant part of the total inertia of the dataset.

The two retained factors that better described the linear correlation coefficient between each of the principal components and the variables being studied were selected by examining the factor loading matrix. Indeed, for each variable, the sum of squares of correlations with the two principal components gives us the part of variance of such variables explained by two first factors.

The "reducted structure" obtained from the first two factors was then used to construct the country's typology using the squared Euclidean distance matrix between any pair of countries. An ascending hierarchical classification using the Ward method for class aggregation was applied. This procedure gives us a highlight concerning the number of classes to be considered. Then the cluster analysis by k–means methods was used to obtain a sample partition into four classes, which allowed the identification of sub clusters of countries with a specific epidemiologic profile at the beginning of May 2020.

Considering the level of statistical asymmetry identified for some of the variables under study, a statistical rank analysis was used to obtain a global overview of the countries' rank within each cluster.

Finally, the association between the countries of the main two identified countries clusters and the set of non–pharmaceutical interventions taken by decision–makers along the first period of COVID–19 pandemic was analyzed (measure not taken; measure taken before March 17, 2020; measure taken on March 17, 2020 or after).

Results

Univariate analysis

A univariate statistical analysis of the five quantitative epidemiological variables was performed and can be found in Table 1. This Table shows clearly that even excluding the smallest and the highest observations (respectively, below first quartile and above third quartile), almost all the variables have a very high relative dispersion evaluated here by the Quartile based variance coefficient (QCV=Interquartile Range/Median). The mortality prevalence especially is very heterogeneous across the countries under study, where QCV is approximately 5.5. Complementary, the Box–plot analysis (Figure 1) provides a visualization of the empirical distribution of each variable.

With regard to the total number of COVID–19 cases, Lithuania, Lithuania, Croatia, Poland, Greece, Bulgaria, Hungary, Romania and Slovakia, registered less than 57.8 cases per 100,000 inhabitants (first quartile), while Switzerland, Portugal, Ireland, Italy, Spain, Belgium and the United Kingdom reported, until May 2, more than 238.7 cases per 100,000 inhabitants. Luxembourg appears as an outlier with 625.6 COVID–19 cases per 100,000 inhabitants. This extreme situation may be related to the exceptionally high number of tests carried out to date in a country with a relatively small population.

With regard to the number of COVID–19 deaths, Latvia, Cyprus, Lithuania, Malta, Greece, Bulgaria and Slovenia recorded less than 1.8 deaths per 100,000 inhabitants in contrast to Ireland, Italy, Spain, Belgium, United Kingdom, Netherlands, France and Sweden with more than 21.5 deaths per 100,000 inhabitants. The distribution of this variable is particularly asymmetric, with Spain and Belgium appearing as outlier countries with 53.1 and 67.4 deaths per 100,000 inhabitants, respectively.

Table 1 Univariate analysis of the epidemiologic variables for 32 countries

	N	Mean	SD	Coef. Of Variation	1.° Quartile	Median	3.° Quartile	IQR	IQR/ Med
Cases Per 100 000 (until May 2, 2020)	31	177.2	151.2	0.9	57.8	147.3	238.7	180.9	1.28
Death Per 100 000	31	13.6	17.8	1.3	1.8	4.0	21.5	19.7	5.53
Active cases per 100 000 (until May 4, 2020)	31	82.0	76.1	0.9	25.3	42.7	144.9	119.6	2.82
Serious critical per 100 000 (until May 4, 2020)	31	1.70	1.60	0.9	0.50	1.10	2.4	1.9	1.73
Tests per 100 000 (until May 4, 2020)	31	3024	1863.2	0.6	1382.5	3113	4200.8	2018	0.65
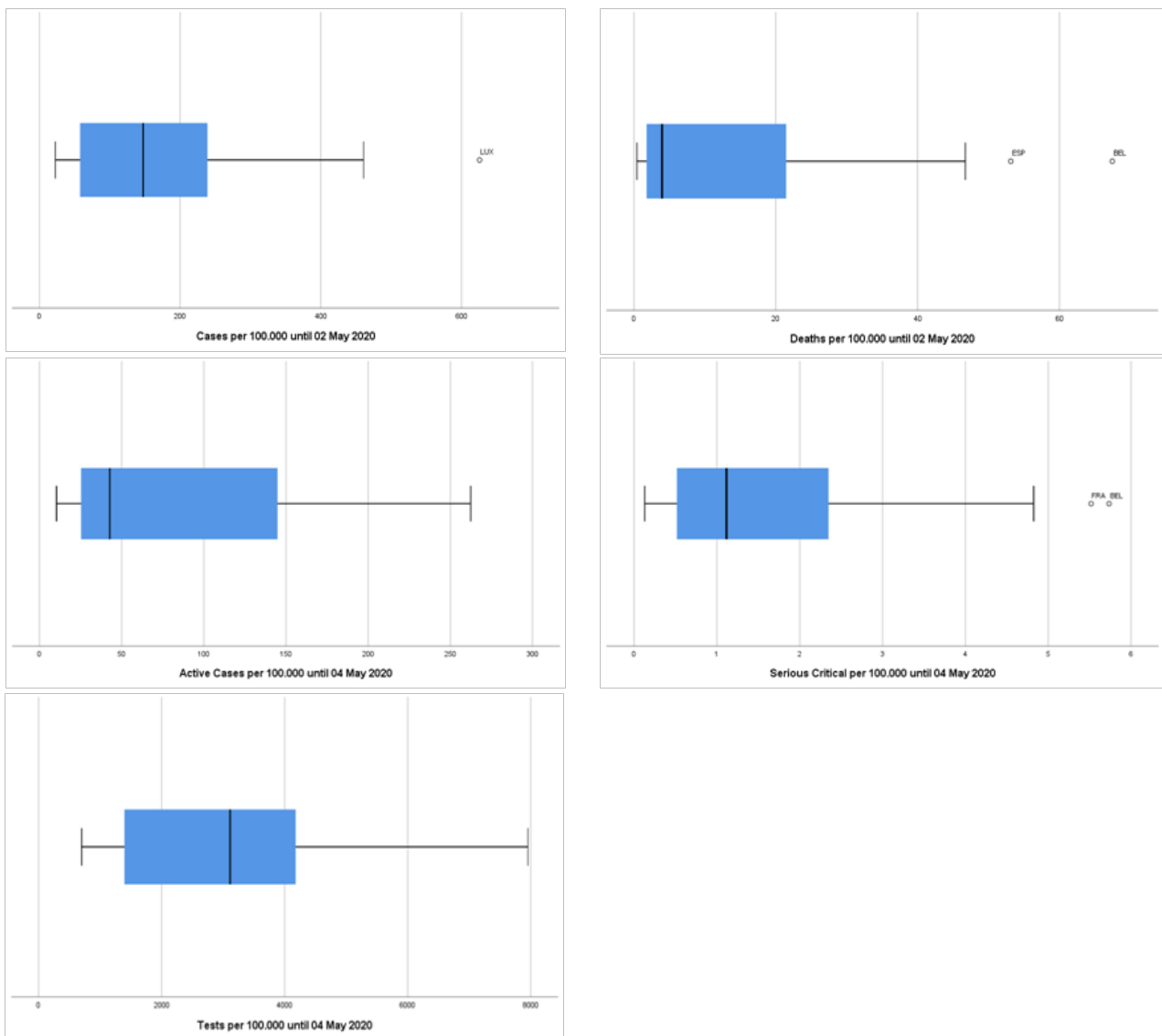
Figure 1 Boxplots for epidemiologic variables.

In the sub universe of reported cases, the number of active cases per 100,000 inhabitants will tend to overestimate the real value. In fact, at the time of this study, some countries had not given priority to the counting of recovered patients, in view of the high pressure healthcare professionals faced to attend the most serious cases and respond to the urgent need to increase the number of tests in pandemic focal areas. Austria, Malta, Croatia, Poland, Greece, Bulgaria, Hungary and Slovakia registered less than 25.3 active cases per 100,000 inhabitants. A particular highlight should be given to the effective response given in Austria. In contrast, Portugal, Ireland, Italy, Spain, Belgium, the United Kingdom, the Netherlands and Sweden showed more than 144.9 active cases per 100,000 inhabitants. In some of these countries the explanation for these values may lie in the higher and growing number of new cases that appeared over the reference period, immediately before our study.

As for the total number of COVID–19 patients in critical condition, it is a particularly sensitive variable, directly related with the logistical capacity of health infrastructures to respond in due time, offering intensive treatments to patients in life threatening conditions. Latvia, Estonia, Malta, Croatia, Poland, Greece and Slovakia registered less than 0.5 critical cases per 100,000 inhabitants. This means that, in these countries, at the time of the study, there may have been no effective pressure on treatment and intensive care services, largely due to the relatively small number of registered COVID–19 cases.

However, in Italy, Spain, Belgium, the Netherlands, France and Sweden, there were more than 2.4 critical cases per 100,000 inhabitants, testing the installed capacity of hospital services despite being inserted in some of the largest economies in Europe. Finally, the Netherlands, Sweden, Croatia, Poland, Greece, Bulgaria, Hungary and Romania had less than 1382.5 per 100,000 inhabitants. In the case of Holland and Sweden, the less strict option in terms of mitigation taken contrasts with the high number of COVID–19 cases registered at the time of our study. Inversely, Portugal, Denmark, Estonia, Israel, Cyprus, Lithuania and Malta have conducted more than 4200.8 tests on May 2 per 100,000 inhabitants.

Bivariate and multivariate analysis

In a second step a bivariate analysis of the data set was applied, previously excluding Iceland and Luxembourg, both outliers regarding the number of tests per 100,000 inhabitants.

The correlation matrix (Table 2) reveals a strong positive correlation between any pair of four epidemiologic variables – cases, deaths, active cases and serious critical – and non–significant linear correlations between these variables and the number of tests. We may thus anticipate two main dimensions discriminating countries: one concerning COVID–19 prevalence and another COVID–19 diagnostic.

Table 2 Linear correlation matrix

	Cases per 100.000 until 02 May 2020	Deaths per 100.000 until 02 May 2020	Active Cases per 100.000 until 04 May 2020	Serious Critical per 100.000 until 04 May 2020	Tests per 100.000 until 04 May 2020
Cases per 100.000	1.0000	0.8392	0.7257	0.7100	0.2687
Deaths per 100.000	0.8392	1.0000	0.7985	0.8712	0.0457
Active Cases per 100.000	0.7257	0.7985	1.0000	0.7060	0.0859
Serious Critical per 100.000	0.7100	0.8712	0.7060	1.0000	–0.0606
Tests per 100.000	0.2687	0.0457	0.0859	–0.0606	1.0000

In a third step a Principal Components Analysis (PCA) was applied to identify a small number of latent variables (principal components) likely to explain a significant part of the total variability of the data. Principal components are latent variables that restore the main part of the initial variability of the data in descending order and exhibit the property of being uncorrelated variables. Therefore, the new matrix constructed by this methodology can provide an interpretation of the data in a reduced dimension space, minimizing the overall loss of variability of the initial data.

In the current situation the mean and variance of the epidemiologic variables are expressed in quite different orders of magnitude, so PCA was based on the correlation matrix. This means that the total variance of the standardized variables is equal to 5, the number of variables. The PCA outputs (Table 3) show that the variance of the first principal component is $\lambda 1 = 3.3$, which means that first principal axis of inertia accounts for 66.8% of the total inertia. On the other hand, the second principal component has a variance of $\lambda 2 = 1.1$ and therefore the second principal axis explains 21.2% of the total inertia.

These two first principal axes of inertia thus explain 88% of the total inertia and hence we consider relatively irrelevant to retain an additional third axis. The examination of the factor loading matrix (Table 4) is crucial for the interpretation of the two retained factors – this Table describes the linear correlation coefficients between each of the principal components and the variables being studied.
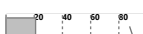
From Table 4 we may conclude that the first factor is a "size factor" positively correlated with all variables. This first axis explains the degree of COVID–19 prevalence across countries, opposing a set of countries where the pandemic had a very low or low prevalence on May 2 to another group of countries were the number of cases, critical cases, death and active cases knew a significant growth along March and April (Figure 2).

In addition, the second principal axis is essentially related to the behavior of countries in what concerns the number of tests per 100,000 inhabitants, opposing countries where such number was above the mean value, to other countries with lower tests numbers.

Table 4 also allows us to represent the epidemiological variables on the first principal plan, in the so–called "correlations circle" (Figure 3), where the coordinates of each epidemiological variable relative to components 1 and 2 are simply the linear correlation coefficients between those variables and the two first principal components. In fact,

all the variables represented here satisfy the condition: r^2 (variable, F1) + r^2 (variable, F2) ≤ 1, where the first member of this inequality corresponds to the part of the variance of each variable under study explained by the first two factors.

Table 3 Principal Components Analysis (PCA): Eigenvalues

Number	Eigenvalue	Percent		Cum Percent
1	3.3411	66.822		66.822
2	1.0616	21.231		88.053
3	0.2941	5.883		93.936
4	0.2163	4.325		98.261
5	0.0869	1.739		100.000

Table 4 Principal Components Analysis (PCA): loadings matrix

	Prin1	Prin2
Cases per 100.000 until 02 May 2020	0.90486	0.19721
Deaths per 100.000 until 02 May 2020	0.96123	–0.08806
Active Cases per 100.000 until 04 May 2020	0.88312	–0.02501
Serious Critical per 100.000 until 04 May 2020	0.89503	–0.22480
Tests per 100.000 until 04 May 2020	0.13190	0.98171
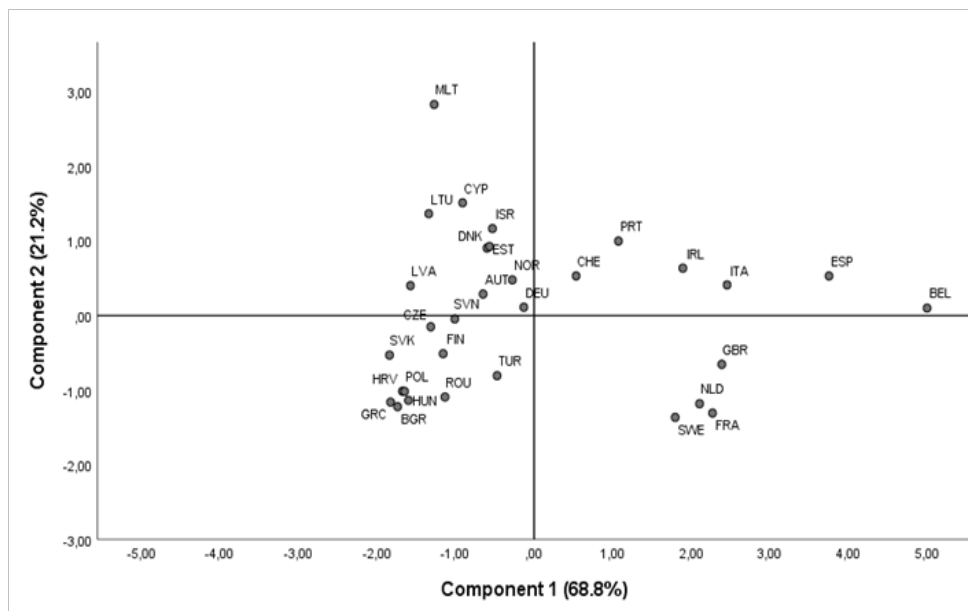
Figure 2 Principal Components Analysis (PCA): Representation of countries on the first principal plan.

Considering our initial aim to characterize statistically the COVID–19 prevalence on May, 2, we used the factorial output explained by the two first factors for the construction of the Euclidean distance matrix between any pair of countries. We applied an ascending hierarchical cluster analysis to this matrix $D_{31 \times 31}$ (Figure 4), taking the

Ward method as aggregation criteria at each iteration of the method.

The dendogram suggested a partition of at least four clusters. This partition was then validated, using K–means non–hierarchical cluster analysis. The results of k–means partitions are suboptimal and

conditioned to the initial choice of number of clusters, here 4 clusters. The final solution will be local optima for the within–group variability (maximizing group homogeneity) and maximizing the between–group separability. The factorial plan 1–2 identifies the cluster partition. Table 5 characterizes each cluster through the "gravity center" of the respective countries.
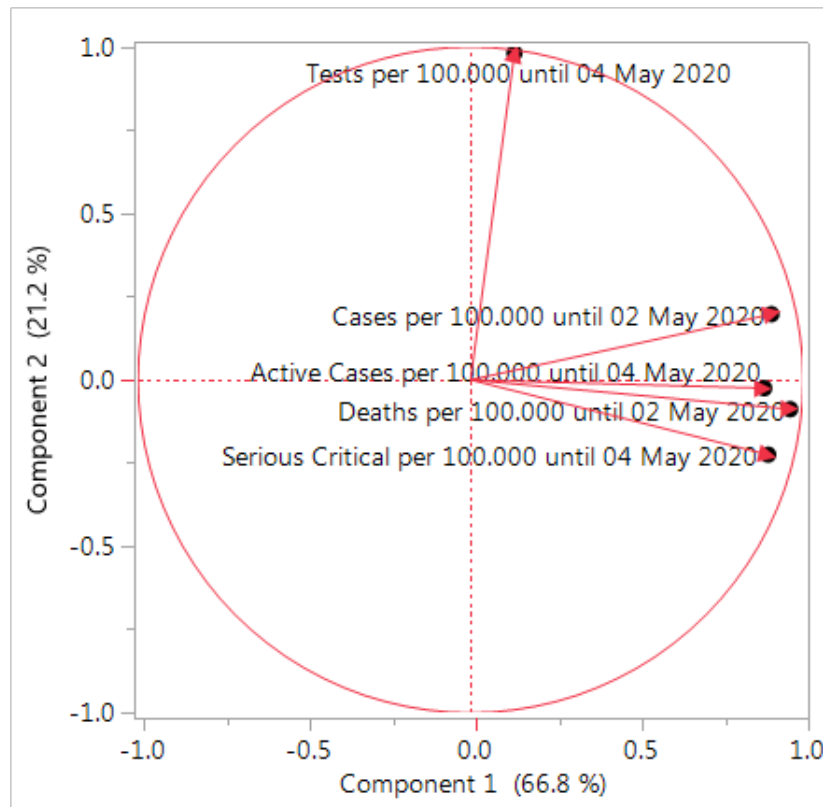
Figure 3 Principal Components Analysis (PCA): Correlation circle.

The first cluster includes the set of 6 countries with a higher total number of cases until May, 2, a mean value of about 376 cases per 100,000 inhabitants (from a minimum of 247 to a maximum of 461). The maximum mean value of deaths also occurs in this cluster (about 37 per 100,000 inhabitants) and the number of tests was the second best among the obtained clusters.

The second cluster, with 4 countries, presents a smaller mean value of total cases (about 226 per 100,000 inhabitants) and a similar mean of deaths and active cases. However, in cluster 2 the serious critical cases were 50% higher than cluster 1 and the number of tests was less than half compared with cluster 1.

Finally, the two other clusters, with 12 and 9 countries, presented a quite smaller number of total cases (a mean of 116 and 55 cases per 100,000 inhabitants, respectively for clusters 3 and 4), the main difference between them being related to the number of tests. Cluster 3 displays a mean of 4,123 tests per 100,000 inhabitants and cluster 4 just 1,155 tests. This significant difference could be explained by the low COVID–19 prevalence in the countries of cluster 4.

A statistical rank analysis was then applied, described in Table 6. Considering the relative asymmetry of most of variables under study, we completed our analysis using a robust approach not conditioned by

biased estimates of mean. From Table 6 it is clear the main difference between clusters 1 and 2 concerns the different priority given to the number of performed tests per 100,000 inhabitants.

The two other clusters describe quite well the rank of countries in what concerns the four epidemiologic variables. Within cluster 3 there are two sub–groups explained again by the number of tests elaborated by respective countries.

Finally, taking the EU countries of clusters 1 and 2, we analyzed eight non–pharmaceutical interventions taken by decision–makers:[15,19]

1. Stay home (enforced)

2. Stay home (General Recommendation)

3. Stay home, Risk Groups

4. Closure secondary and Higher Schools

5. Closure Primary Schools

6. Closure Educational Care

7. Closure Public Services

8. Mass Gathering cancelations
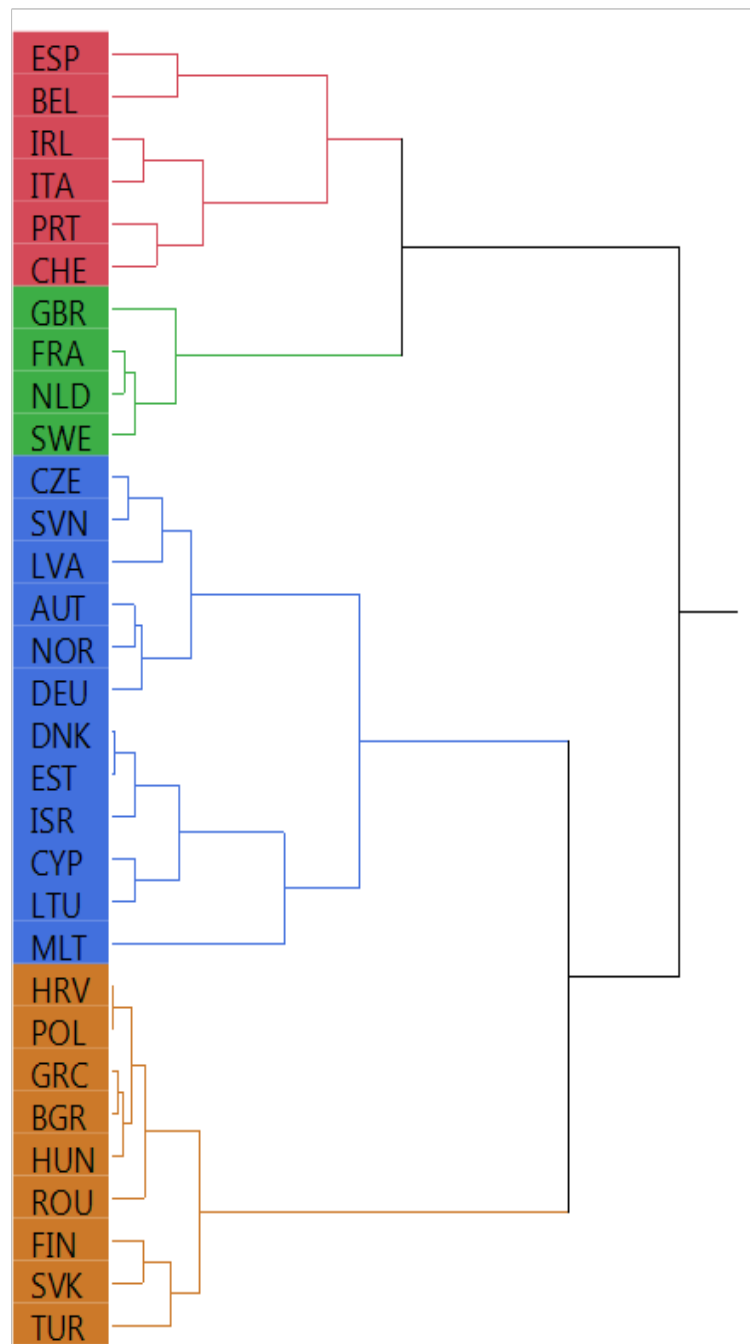
Figure 4 Ascending Hierarchical Cluster Analysis – Dendrogram.

Clusters 1 and 2 included nine countries where the effects of the pandemic took a more worrying proportion in terms of health but also in economic and social terms. Table 7 indicates the measures implemented in each of the nine countries, as well as the date on which they were undertaken, concerning the period between February 25 and April 25.

Five main results may be observed from Table 7:

• Sweden's strategy towards COVID–19 was quite different from that of the other countries: just "Stay home – Risk Groups", "Closure of Secondary and Higher Schools" and "Mass Gathering cancelations" were considered and implemented later than other countries;

• The Belgian and French strategies did not recommend staying at home, even for groups at risk;

• The Italian strategy did not recommend staying home at an early stage of the pandemic and only later was decided as an enforced measure;

• United Kingdom implemented all measures, but at a later stage in comparison with the other countries;

• The "Mass Gathering cancelations" measure was implemented at an early stage of COVID–19 in Ireland, France, Portugal and Italy. In contrast, the UK only implemented this measure a month later.
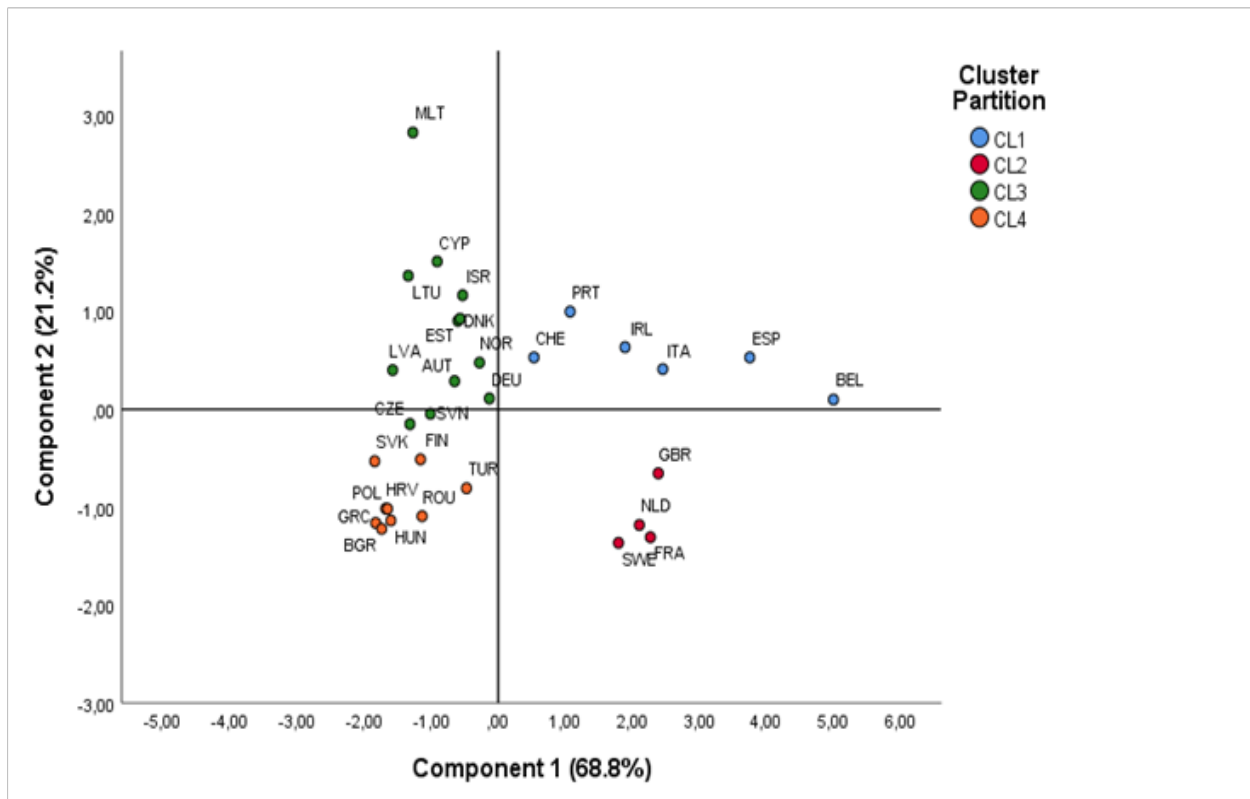
Figure 5 Principal Components Analysis (PCA): Representation of each cluster on the first factorial principal plan.

Table 5 Characterization of each cluster of countries

| | | **Cluster Partition** | | | |
		CL1	**CL2**	**CL3**	**CL4**
Cases per 100.000 until 02 May 2020	Mean	376,13	225,87	116,06	54,76
	Standard Deviation	79,31	31,17	54,59	42,00
	Minimum	246,56	194,34	45,16	22,61
	Maximum	460,61	266,89	194,99	148,68
	Valid N	6	4	12	9
	Median	388,53	221,13	112,41	34,51
	Percentile 25	343,25	202,84	70,71	25,76
	Percentile 75	429,27	248,91	168,08	64,53
Deaths per 100.000 until 02 May 2020	Mean	36,66	33,13	3,71	2,38
	Standard Deviation	22,57	7,15	2,59	1,41
	Minimum	9,79	26,05	,83	,42
	Maximum	67,44	41,38	7,93	3,96
	Valid N	6	4	12	9
	Median	36,39	32,56	3,19	1,83
	Percentile 25	16,84	27,22	1,65	1,31
	Percentile 75	53,13	39,04	5,58	3,82

Citation: 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. DOI: 10.15406/bbij.2020.09.00309

Table Continued

		Cluster Partition			
		CL1	**CL2**	**CL3**	**CL4**
Active Cases per 100.000 until 04 May 2020	Mean	163,72	185,79	50,77	26,04
	Standard Deviation	77,48	47,50	38,25	18,05
	Minimum	35,19	138,67	15,92	10,37
	Maximum	262,34	242,91	144,10	68,07
	Valid N	6	4	12	9
	Median	158,96	180,79	35,71	21,03
	Percentile 25	145,61	147,30	26,98	13,68
	Percentile 75	221,26	224,29	60,10	28,76
Serious Critical per 100.000 until 04 May 2020	Mean	2,99	4,07	,84	,68
	Standard Deviation	1,83	1,32	,59	,50
	Minimum	1,39	2,34	,21	,13
	Maximum	5,73	5,52	2,35	1,68
	Valid N	6	4	12	9
	Median	2,18	4,22	,70	,53
	Percentile 25	1,66	3,15	,48	,37
	Percentile 75	4,82	4,99	1,12	,89
Tests per 100.000 until 04 May 2020	Mean	3801,16	1531,90	4122,54	1155,01
	Standard Deviation	401,40	334,57	1420,13	438,03
	Minimum	3320,01	1173,50	2461,47	699,65
	Maximum	4376,68	1942,56	7487,44	1875,66
	Valid N	6	4	12	9
	Median	3742,30	1505,77	3826,23	1010,58
	Percentile 25	3489,79	1271,30	3112,90	859,45
	Percentile 75	4135,92	1792,50	4857,00	1422,67

Table 6 Statistical rank analysis by countries and associated clusters

Cluster	Countries/ VAR	Cases Per 100 000	Death Per 100 000	Active cases per 100 000	Serious critical per 100 000	Tests per 100 000
1	CHE	4	3	2	3	3
	PRT	4	3	4	3	4
	IRL	4	4	4	3	3
	ITA	4	4	4	4	3
	ESP	4	4	4	4	3
	BELG	4	4	4	4	3
	UK	4	4	4	3	2
2	NL	3	4	4	4	1
	FR	3	4	3	4	2
	SE	3	4	4	4	1

Table Continued

Cluster	Countries/ VAR	Cases Per 100 000	Death Per 100 000	Active cases per 100 000	Serious critical per 100 000	Tests per 100 000
	CZE	2	2	2	2	2
	SVN	2	2	3	2	2
	LVA	1	1	2	1	3
	AUT	3	3	1	3	3
	NOR	2	2	3	2	3
3	**DEU**	3	3	2	3	2
	DNK	3	3	2	2	4
	EST	2	2	3	1	4
	ISR	3	2	3	2	4
	CYP	2	1	3	3	4
	LTU	1	1	2	2	4
	MLT	2	1	1	1	4
	HVR	1	2	1	1	1
	POL	1	2	1	1	1
	GRC	1	1	1	1	1
	BGR	1	1	1	2	1
4	**HUN**	1	2	1	2	1
	ROU	2	2	2	3	1
	FIN	2	2	2	2	2
	SVK	1	1	1	1	2
	TUR	3	2	3	3	2

1: Below first quartile

2: Between first quartile and median

3: Between median and 3.° quartile

4: Above 4.° quartile

Table 7 List of non–pharmaceutical governmental measures against COVID–19 (until May, 2)

		Recomend Group	Risk Group	Closure Second Higher	Closure Second Higher	Closure	Closure	Mass Gattering Cancellations
PRT	NO	at 17/3	at 10/3	at 12/3	at 12/3	at 12/3	at 12/3	at 04/3
IRL	at 26/3	from 13/3 to 25/3	at 13/3	at 13/3	at 13/3	at 13/3	at 13/3	at 25/2
ITA	at 12/3	NO	at 04/3	at 04/3	at 04/3	at 04/3	at 12/3	at 04/3
ESP	at 15/3	from 08/3 to 14/3	at 12/3	at 12/3	at 12/3	at 12/3	at 14/3	at 14/3

Table Continued

		Recomend Group	Risk Group	Closure Second Higher	Closure Second Higher	Closure	Closure	Mass Gattering Cancellations
BELL	at 20/3	NO	NO	at 13/3	at 13/3	NO	at 13/3	at 10/3
UK	at 24/3	from 17/3 to 23/3	at 17/3	at 22/3	at 22/3	at 22/3	at 17/3	at 24/3
NL	at 25/3	from 17/3 to 22/3	at 13/3	at 12/3	at 17/3	at 17/3	at 12/3	at 12/3
FR	at 19/3	NO	NO	at 17/3	at 17/3	at 17/3	at 17/3	at 28/2
SE	NO	NO	at 20/3	at 18/3	NO	NO	NO	at 12/3

Discussion

The objective of this cross–sectional study was to describe the COVID–19 prevalence at the beginning of May 2020, in European Union Countries also belonging other European OECD countries.

The epidemiological variables under study were: total number of COVID–19 cases, total number of deaths, total number of active cases, total number of critical serious cases and total number of tests. The first four variables were positively correlated and uncorrelated with the number of tests. In order to obtain a typology of countries that explained their COVID–19 prevalence, four clusters were chosen. Clusters 1 and 2 include countries where the epidemic has reached a higher impact than the global average in the first four variables under study. However, among these nine countries, Switzerland and Portugal presented better results in what concerned the numbers of deaths and of critical serious cases per 100,000 inhabitants. We also emphasize the fact that clusters 1 and 2 are distinguished by the number of tests performed. In fact, the number of tests in cluster 2 was significantly smaller, especially in Sweden and the Netherlands with values below the global first quartile. For the cluster 1, all the countries presented a number of tests above the global median with emphasis on Portugal, where this value was above the third global quartile. The analysis of the third cluster revealed a lower COVID–19 prevalence, although seven of the twelve countries in this cluster had at least one of the four epidemiologic variables with values above the global median. Cluster 3 featured six countries where the number of tests exceeded the third global quartile. Cluster 4 included a set of nine countries with heterogeneous geographies where the COVID–19 prevalence was particularly low, generally below the global median or even below the first quartile. Turkey was an exception with the number of cases namely active cases per 100,000 inhabitants and the number of serious critical cases above the global median. Due to the asymmetry of the empirical distribution of some of the epidemiological variables, and the consequent bias in the respective means, we validated the previous results using the statistical rank analysis.

Regarding our second aim, we established some associations between the measures implemented by the Governments in the first phase of the pandemic and the COVID–19 prevalence data over the period until early May 2020.

This investigation covered just EU countries belonging to the two first clusters, the countries hardest hit by COVID–19. The eight selected measures preventing the degree of contamination of this viral epidemic were: stay at home (enforced), stay home recommendation, stay home for risk group, closure of secondary and higher schools, closure of primary schools, closure of educational care day, closure of public services and mass gathering cancellations.

We have associated to each measure three modalities: not implemented, implemented at an early date (before March 17, 2020), and implemented at a later date (on or after March 17, 2020). Global behaviors of the EU countries of clusters 1 and 2 were represented by a disjunctive table. Our approach generated a cluster analysis which reproduced the two classes of countries previously obtained, with the exception of Sweden, which constituted an isolated group in views of this country's completely different strategy in combating COVID–19.

Timings of the implementation of non–pharmaceutical interventions may thus have had an impact of the structure of the obtained country clusters. This impact may be assessed in terms of prevalence indicators, but also in economic terms, and both are relevant in the future to identify the best approaches.[19–21]

Future studies will further analyse the following cycles, as less restrictive measures will be undertaken regarding the mobility of populations, economic activity and more advanced deconfinement phases.

130

Conclusions

In the present paper four classes were identified to obtain a typology of countries that would describe their COVID–19 prevalence. Essentially, we identify a set of ten countries where such prevalence generates a real concern, testing throughout the period under study the capacity of health services. Fortunately, in same countries (like Switzerland and Portugal) the number of serious critical cases did not exceed the health logistical available, although in Portugal patients' access to non–COVID–19 health services may have been hindered.[22] There is some evidence that non–pharmaceutical measures taken from February, 25 until April, 25 had the desire impact of controlling the epidemic within reasonable limits, but we know that the key aims of these measures will be to reduce the effective reproduction number R_t, a quite crucial epidemiological quantity representing the average number of infection generated at time t, by each injected case over the course of their infections.

However, data referenced of our present study do not access some potentially relevant equity issues, like the impact of Swedish strategy in terms of COVID–19 prevalence over some risk level groups like the elderly.

Our next studies will try to identify which interventions are necessary to maintain control over epidemiological variables, while also manage the trade–off regarding their impact on economic, financial and social costs.

Acknowledgments

None.

Conflicts of interest

The authors report no conflicts of interest in this work.

References

1. Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. 2020;382:1708–1720.

2. Mole B. Don't Panic: The comprehensive Ars Technica guide to the coronavirus. 2020.

3. Lauer SA, Grantz KH, Bi Q, et al. The Incubation Period of Coronavirus Disease 2019 (COVID–19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med*. 2020;172(9):577–582.

4. 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;395:1054–1062.

5. Levi M, Scully M. How I treat disseminated intravascular coagulation. *Blood*. 2018;131(8):845–854.

6. 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.

7. 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.

8. Giannis D, Ziogas IA, Gianni PJ. Coagulation disorders in coronavirus infected patients: COVID–19, SARS–CoV–1, MERS–CoV and lessons from the past. *J Clin Virol*. 2020;127:104362.

9. Wu Z, McGoogan 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;323:1239–1242.

10. Mehta P, McAuley DF, Brown M, et al. COVID–19: consider cytokine storm syndromes and immunosuppression. *Lancet*. 2020;395(10229):1033–1034.

11. Klok FA, Kruip M, van der Meer NJM, et al. Incidence of thrombotic complications in critically ill ICU patients with COVID–19. *Thromb Res*. 2020;191:145–147.

12. Cascella M, Rajnik M, Cuomo A, et al. Features, Evaluation and Treatment Coronavirus (COVID–19). 2020 May 18. In: StatPearls [Internet]. Treasure Island (FL): *StatPearls Publishing*; 2020.

13. 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(5):1094–1099.

14. 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 Pract*. 2020.1–11.

15. European Centre for Disease Prevention and Control. Rapid Risk Assessment: Coronavirus disease 2019 (COVID–19) in the EU/EEA and the UK– ninth update. 2020.

16. Worldometers.info. COVID–19 Coronavirus Pandemic, 2020.

17. European Centre for Disease Prevention and Control. COVID–19 Pandemic, 2020.

18. Flaxman S, Mishra S, Gandy A. et al. Estimating the effects of non–pharmaceutical interventions on COVID–19 in Europe. *Nature*. 2020.

19. European Centre for Disease Prevention and Control. Coronavirus disease 2019 (COVID–19) in the EU/EEA and the UK – tenth update, 11 June 2020, ECDC. 2020.

20. Ferguson N. Report 9: Impact of non–pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand – Working paper, Imperial College COVID–19 Response Team, London, UK, 16 March 2020.

21. Haushofer J, Metcalf CJE. Which interventions work best in a pandemic?, Science. 2020.

22. Nogueira PJ, Araújo–Nobre M, Nicola PJ, et al. Excess mortality estimation during the COVID–19 pandemic: preliminary data from Portugal, *Acta Médica Portuguesa*. 2020;33(6):376–383.