BMJ Open

Using hierarchical clustering analysis to evaluate COVID-19 pandemic preparedness and performance in 180 countries in 2020

Banafsheh Sadeghi [1], Rex C Y Cheung,[2] Meagan Hanbury[1]

Check for updates

[1]Health Bridge Analytics, Davis, California, USA
[2]Department of Decision Sciences, San Francisco State University, San Francisco, California, USA

Correspondence to
Dr Rex C Y Cheung;
rexcheung@sfsu.edu

ABSTRACT

Objective To rank and score 180 countries according to COVID-19 cases and fatality in 2020 and compare the results to existing pandemic vulnerability prediction models and results generated by standard epidemiological scoring techniques.

Setting One hundred and eighty countries' patients with COVID-19 and fatality data representing the healthcare system preparedness and performance in combating the pandemic in 2020.

Design Using the retrospective daily COVID-19 data in 2020 broken into 24 half-month periods, we applied unsupervised machine learning techniques, in particular, hierarchical clustering analysis to cluster countries into five groups within each period according to their cumulative COVID-19 fatality per day over the year and cumulative COVID-19 cases per million population per day over the half-month period. We used the average of the period scores to assign countries' final scores for each measure.

Primary outcome The primary outcomes are the COVID-19 cases and fatality grades in 2020.

Results The United Arab Emirates and the USA with F in COVID-19 cases, achieved A or B in the fatality scores. Belgium and Sweden ranked F in both scores. Although no African country ranked F for COVID-19 cases, several African countries such as Gambia and Liberia had F for fatality scores. More developing countries ranked D and F in fatality than in COVID-19 case rankings. The classic epidemiological measures such as averages and rates have a relatively good correlation with our methodology, but past predictions failed to forecast the COVID-19 countries' preparedness.

Conclusion COVID-19 fatality can be a good proxy for countries' resources and system's resilience in managing the pandemic. These findings suggest that countries' economic and sociopolitical factors may behave in a more complex way as were believed. To explore these complex epidemiological associations, models can benefit enormously by taking advantage of methods developed in computer science and machine learning.

INTRODUCTION

After centuries of confronting international public health emergencies, world health authorities have accumulated an enormous amount of experience and knowledge

Strengths and limitations of this study

► Applying unsupervised machine learning method to reveal hidden patterns in COVID-19 data.
► Comparing the results with classical epidemiological methods used in other published studies.
► Comparing the results among countries within similar area and population size groups.
► Potential misclassification of COVID-19 cases and fatality numbers due to lack of a unified definition, case finding method and diagnostic procedure.
► Lack of any validation study against a gold standard in preparedness and disease management measures.

regarding countries' preparedness for pandemics. The potential transnational spread of infectious diseases makes it imperative that all countries are prepared to an adequate level. Preparedness is a complex construct to measure due to its multifactorial predictors at every level of a country's structure and functionality. Previous efforts to assess countries' pandemic preparedness, such as the WHO Joint External Evaluation (WHO-JEE)[1] and the RAND pandemic preparedness project,[2] were based on available data reflecting countries' attributes and expert opinions.

In the 58th World Health Assembly in 2005, the assembly adopted the International Health Regulations resulting in the development of the JEE taskforce to assess countries' preparedness for infectious outbreaks. The JEE developed a questionnaire tool targeting three capacity domains including prevention, detection and rapid response to public health threats.[1] The JEE evaluation emphasised the importance of countries' voluntary participation, a multisectoral approach, evaluation by both external experts and countries' internal self-assessment, data transparency and information sharing. Countries' preparedness

improvement was envisioned as a continuous effort towards preset international targets through recurring assessments every 5 years. Each JEE capacity domain has an indicator score ranging from '1' (implementation has not occurred) to '5' (implementation has occurred, been tested, reviewed and exercised and that the country has a sustainable level of capability for the indicator). For each domain, a country receives a single score based on their current capacity level.

In 2016, the International Security and Defense Policy Center of the RAND National Defense Research Institute created the Infectious Disease Vulnerability Index, a data-driven tool incorporating seven domains that influence countries' vulnerability to outbreaks. Domains include demography, healthcare, public health, disease dynamics, domestic policies, international policies and economy.[2] The RAND research team used available data, including WHO and World Bank databases, along with experts' opinion to calculate the Indexed Vulnerability Scores (IVS) for each country. Scores range between 0 (most vulnerable to infectious disease outbreaks) and 1 (most resilient). One hundred and eighty-eight countries were evaluated and ranked by overall IVS and the seven domain scores. In an empirical analysis of the vulnerability algorithm using Zika (2015–2016) and Ebola (2014–2015), the final report concluded that the IVS tool did not predict success or failure in outbreak response.[3] From December 2019 when COVID-19 was first identified in Wuhan, China, to 11 March 2020, when the WHO declared a pandemic, there were 118 000 confirmed COVID-19 cases in 114 countries.[4] By the end of 2020, there were 83 872 473 confirmed cases and 1 826 809 deaths, worldwide.[5] World health authorities provided guidelines and leadership to contain the virus, yet countries' responses to the virus varied. Containing an outbreak and adopting strategies to slow the spread of a virus requires a high level of public health emergency preparedness and coordination. The total number of cases relative to the population in a country can serve as a proxy for the country's preparedness and a performance measure reflecting how effectively the systems function. Fatality, however, can be a proxy measure for crisis management with respect to healthcare system capacity and flexibility to offer targeted care.

This paper applies hierarchical clustering analysis (HCA) to rank and score 180 countries according to COVID-19 cases and fatality in 2020. We also compare our results to WHO-JEE and RAND predictions and outcomes generated by standard epidemiological scoring techniques.

METHODS
Data
In this retrospective observational study, we used three data sources. The number of countries with full data varied across sources.

1. COVID-19 data: we used the Oxford COVID-19 Government Response Tracker to obtain COVID-19 cases per day and fatality for 180 countries from 1 January to 31 December 2020.[6]
2. RAND data: we used the overall IVS and seven domain scores for 195 countries.[3]
3. WHO-JEE data: we used the Ready Score and four subscores (prevent, detect, respond, and other) for 98 countries.[7]

Patient and public involvement
This study was solely by using publicly available data, and there was no patient or public involvement in any part of the study or final report.

Clustering analysis
In machine learning, most algorithms can be broadly divided into two categories: supervised and unsupervised algorithms. While supervised algorithms aim to learn the relationship between input data and the response variable, unsupervised algorithms provide a way to uncover hidden structure within a dataset. Cluster analysis is a set of unsupervised algorithms that aim to group n observations into K groups (clusters), where intragroup observations are similar and intergroup observations are more distinct. There are several popular clustering algorithms, including K-Means and HCA. Our analysis uses HCA to cluster similar countries into groups. We used agglomerative hierarchical clustering as it is more popular and intuitive than divisive hierarchical clustering. We broke the year into 24 half-month periods and clustered countries into K=5 groups within each period according to their cumulative COVID-19 fatality per day over the year and cumulative COVID-19 cases per million population per day over the half-month period. By clustering on a half-month basis, the scoring process accounts for time-related factors shared by countries. In each period, each observation represents a country with roughly 15 measurements.

Country categories
Country size influences population-adjusted rankings of COVID-19 outcomes. To provide a fair comparison, we divided countries into three categories based on population and land area. We did this via hierarchical clustering with the Euclidean measure between observations' distance and the Ward criterion for forming the clusters.[8] The variables used were the z-score standardised logarithmic transformation of population and land area (in km^2). Four clusters initially categorised countries into micro, small, medium and large sizes. The micro and small countries were combined to form a larger group of small countries. The resulting categories were small (n=58), medium (n=60) and large (n=62). Clustering of cumulative fatality and cumulative cases per million population was performed separately on each subcategory.

Half-month cases per million population
We calculated the cumulative COVID-19 new cases per million population for each day in half-month periods.

We used cumulative case counts to have a picture of changes in the overall case rate trend within a 2-week period (half a month) and avoid the random variation of the daily new case count. This measure is set back to zero at the beginning of each half-month period. As we lived through the pandemic, more information was provided, and various policies were implemented, hence resetting the case count helps explore the impact of time on each country. By using the per million population unit in the denominator, incidence numbers were comparable across countries with different population sizes. Cluster analysis was performed as described previously. Each country was given a score based on the assigned cluster, with a score of 1 presenting a cluster with the highest and 5 the lowest average cases per million population. Due to the low number of COVID-19 cases in the first half of January, these data were omitted during the clustering process. The analysis included 23 half-month period clustering results for cases per million population (mid-January to December).

Half-month cumulative fatality

We calculated the cumulative fatality per day for each country by taking the ratio between the cumulative COVID-19 deaths and cumulative COVID-19 cases on each day. Clustering and scoring were performed as described previously, where a score of 1 indicating the group with the highest fatality. Due to the low number of COVID-19 deaths in January, we omitted these data when performing the clustering analysis. The resulting analysis yielded 22 half-month period clusters for fatality (from February to December). It is conventional to standardise the data (by subtracting each measure from the mean and dividing by the SD) before performing clustering to avoid unequal contribution due to uneven scales. Even though uneven scales in our variables are unlikely, we performed scaling as a sanity check. We conducted the same analysis with scaled data and discovered no significant difference, hence the unscaled version is presented.

In contrast with the cumulative cases, the cumulative fatality was not reset to zero at the beginning of each half-month period. Deaths are independent that implies that an individual's death does not change the risk of other cases to die. The case counts, however, are dependent measures: the higher number of new cases results in the higher likelihood of contracting the disease. Consequently setting the cumulative case counts back to zero enables us to detect countries' other time-dependent factors such as shifting in public health policies.

Final scores, quarterly scores and grades

Our primary analysis yielded 22 fatality and 23 case scores for each country. We used the average of the period scores to assign countries final scores for each measure. We normalised the final scores such that all scores are within the range of 0–1. For each final score X_i, the normalised version is $(X_i - X_{min}) / (X_{max} - X_{min})$, where X_{min} and X_{max} are the minimum and maximum final scores for fatality

and cases. Normalisation was done separately for each of the three country size categories, and normalised final scores are presented.

To assess the evolution of clusters from a broader perspective, we repeated our primary analysis using 3-month periods and assigned quarterly scores. No averaging was done for this subanalysis.

We translate results into A, B, C, D and F grades, where A is excellent and F is the worst grade in performance. Quarterly grades are equivalent to the clusters, while overall 2020 grades are based on five equal segments of the normalised final scores (ie, scores from 0 to 0.19 are F, from 0.20 to 0.39 are D and so on.)

Comparison points

For comparison with standard epidemiological methods, we present both case and fatality 'naïve averages' as well as 31 December values. Naïve averages were calculated by first averaging COVID-19 cases per million population and cumulative fatality values for 22 and 23 half-month periods, respectively, for each country. Then half-month averages were averaged into a final value for each country and measure combination. December 31 values reflect the total numbers of COVID-19 cases per million population and cumulative fatality over the year.

Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of our research.

RESULTS

Figure 1 presents two sets of world maps, one for COVID-19 case final scores and one for COVID-19 fatality final scores. Each set is broken down into three groups of small, medium and large countries. Table 1 presents the list of countries with C, D and F grades for COVID-19 cases, both by quarter and for 2020. Among the small countries, Andorra, Bahrain, Luxembourg, Panama, Qatar and San Marino received an F for 2020. Among the medium countries, Belgium, Israel, Oman, Switzerland, Sweden and United Arab Emirates were graded F for 2020. Finally, among the large countries, Spain and the USA received an F for 2020. Looking at the quarterly breakdown, all the countries with an overall F grade had C, D or F in at least one quarter.

Table 2 presents the list of countries with C, D and F grades of COVID-19 fatality, both by quarter and for 2020. Among the small countries, Bermuda, Gambia, Guyana, Ireland, San Marino and Puerto Rico ranked F. In the medium countries, Belgium, Ecuador, Hungry, Liberia, Netherlands and Sweden ranked F. Among the large countries, only Yemen received F for 2020 and in most quarters (online supplemental file provides the score ranks for the 180 countries).

Figure 2 shows the scatter plots of the RAND, WHO-JEE, naïve average and 31 December score distributions against HCA scores. The pandemic preparedness scores
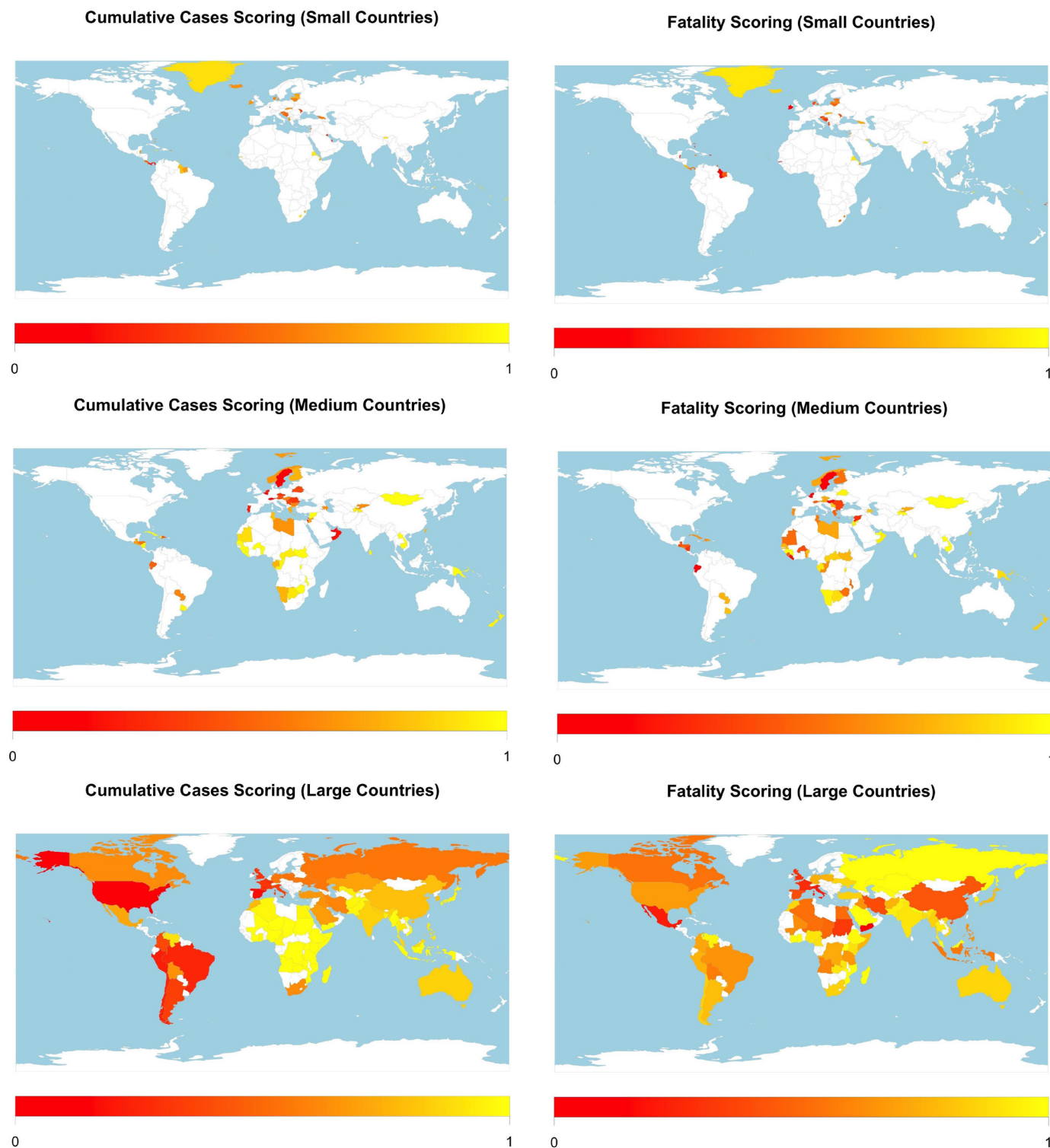
Figure 1 Hierarchical cluster analysis scores based on country size. 0 reflects worst scores, 1 reflects best scores.

calculated by RAND and WHO-JEE had no relationship with COVID-19 cases or fatality measures. In correlation analysis (details in web reference), there was consistently an inverse relationship between the WHO-JEE and RAND scores with the HCA scores. The closest correlation between RAND and HCA was −0.10 for fatality in large countries. For WHO-JEE, the closest correlation was 0.05 for fatality in small countries. Among the RAND

subcategory domains, disease dynamic had the highest correlation with the HCA scores, with a value of 0.10 for fatality in small countries. For JEE, the closest correlation was 0.17 for fatality for small countries in the detect subcategory. Both naïve averages and 31 December values have closer alignment with HCA scores. For naïve averages, correlations range from 0.88 to 0.97 for cases and 0.92 to 0.97 for fatality. For 31 December values, correlations

Table 1 Quarterly and the year C, D and F grades of COVID-19 cases in 2020

Q1	Q2	Q3	Q4	2020 score	
Small countries					
San Marino	Qatar	Aruba	Andorra	Bahrain	0
Andorra	San Marino	Bahrain	Luxembourg	Andorra	0.082
Luxembourg	Andorra	Panama	Slovenia	Luxembourg	0.143
Iceland	Bahrain	Qatar	Georgia	San Marino	0.143
Faeroe Islands	Ireland	Kuwait	Guam	Panama	0.163
	Kuwait	Costa Rica	Bosnia and Herzegovina	Qatar	0.163
	Singapore	US Virgin Islands	San Marino	Kuwait	0.265
			Croatia	Moldova	0.347
			Lithuania	Aruba	0.429
			Slovak Republic	Bosnia and Herzegovina	0.469
				Costa Rica	0.469
				Guam	0.469
				Slovenia	0.49
				Croatia	0.51
				Cape Verde	0.531
				Georgia	0.551
				Palestine	0.551
				Singapore	0.551
				Lithuania	0.571
				Malta	0.571
				Denmark	0.592
				Puerto Rico	0.592
Medium countries					
Switzerland	Belgium	Israel	Belgium	Sweden	0
Belgium	Sweden	Oman	Czech Republic	Israel	0.022
Austria	United Arab Emirates	Dominican Republic	Switzerland	Belgium	0.044
Norway	Belarus	Kyrgyz Republic	Netherlands	United Arab Emirates	0.089
Netherlands	Oman		Serbia	Switzerland	0.156
Sweden	Switzerland		Austria	Oman	0.178
Portugal	Netherlands		Sweden	Portugal	0.2
Israel	Portugal		Portugal	Netherlands	0.222
	Israel		Romania	Czech Republic	0.311
	Serbia		Bulgaria	Romania	0.311
	Ecuador		Hungary	Serbia	0.356
			Jordan	Austria	0.378
				Belarus	0.378
				Dominican Republic	0.422
				Ecuador	0.467
				Honduras	0.533
Large countries					
Italy	Chile	USA	USA	USA	0

Continued

Table 1 Continued

	Q1	Q2	Q3	Q4	2020 score	
	Spain	USA	Peru	France	Spain	0.17
	USA	Peru	Brazil	Italy	Peru	0.226
	France	Italy	Argentina	Poland	Brazil	0.245
	UK	Spain	Colombia	Argentina	Chile	0.245
	Germany	France	Chile	Spain	France	0.302
	Iran	UK	Bolivia	Ukraine	UK	0.302
		Turkey	Kazakhstan	UK	Argentina	0.34
		Canada	South Africa	Brazil	Italy	0.34
		Brazil	Spain	Colombia	Colombia	0.377
		Russia	Iraq	Russia	Poland	0.528
		Saudi Arabia		Turkey	Germany	0.547
				Germany	Russia	0.547
					Ukraine	0.547
Grade F						
Grade D						
Grade C						

range from 0.88 to 0.97 and 0.64 to 0.80 for cases and fatality, respectively (the HCA values were inversed to follow the same direction as averages).

DISCUSSION

With 83 872 473 confirmed cases and 1 826 809 deaths due to COVID-19 in 2020, there is a growing need to examine countries' performance in controlling the virus and managing infected cases. The assessment of health policies and preventive interventions needs to be based on a robust performance evaluation methodology. In classic epidemiology, different measures of normalised counts such as averages, and rates are common methods of scoring performances. These measures either present the central tendency (ie, arithmetic mean, mode or median) or the peripheral distribution (ie, SD or percentiles) and fall short in capturing the behaviour of data over time and relative to other data points. In terms of the COVID-19 pandemic and ranking countries' performance, there are several epidemiological characteristics that need to be considered in evaluation studies. Knowledge about COVID-19 transmission and pathophysiology was an emerging phenomenon and as a result guidelines and policies were updated continuously over time. Some countries prioritised their economy over complete shutdown and chose herd immunity over transmission prevention, whereas others enforced more restrictive policies, such as sheltering in place and closing non-essential businesses. The waves of COVID-19 cases and deaths varied among continents, regions and countries, while factors such as countries' proximities, international transportation routes and economic treaties influenced COVID-19 transmission and violated the data independency assumption among countries. The complex situation of the COVID-19 pandemic calls for more sophisticated methods in performance evaluation than averages and rates. An ideal performance evaluation method will assess the changes of COVID-19 cases and fatality over time and relative to other countries. In a paper published by Jamison et al[9] the authors proposed the use of COVID-19 cases and death doubling time to avoid the inadequacy of cumulative deaths per million population. Although doubling time may provide a more informative measure, it still falls short in providing a dynamic approach to compare countries with their own performance and others' over time.

Our methodology enabled us to account for countries' performance over time by analysing COVID-19 cases and fatality in half-month periods, while HCA allowed us to construct hierarchies of similarly performing countries. To our knowledge, no other reports present 2020 world performance based on cluster analysis in a dynamic setting. However, there are several studies that use basic cluster analysis on COVID-19 data. Papers from India[10 11] Brazil[12] and Indonesia[13] assess number of confirmed cases and deaths on local scales. Nascimento[12] was the only study to use additional factors in their clustering, including population, land area, gross domestic product, population density and the Human Development Index. They performed clustering analysis on a weekly basis from 3 April to 8 August, while the other local level studies performed clustering using single day data. In a study of COVID-19 cases of the 50

Table 2 Quarterly and the year C, D and F grades of COVID-19 fatality in 2020

Q1	Q2	Q3	Q4	2020 Scores	
Small Countries					
Guyana	Guyana	San Marino	San Marino	San Marino	0
Gambia	San Marino	Andorra	Bermuda	Guyana	0.042
San Marino	Bahamas	Bermuda	Fiji	Ireland	0.127
	Belize	Barbados	Barbados	Bermuda	0.141
	Suriname	Ireland	Ireland	Gambia	0.183
	Andorra	Slovenia	Guyana	Puerto Rico	0.197
	Bosnia and Herzegovina	Guyana	Bosnia and Herzegovina	Barbados	0.211
	Bermuda	Bahamas	Gambia	Bahamas	0.239
	Barbados	Belize	Slovenia	Bosnia and Herzegovina	0.282
	Denmark	Denmark	Bahamas	Slovenia	0.296
	Ireland	Puerto Rico	Belize	Belize	0.31
	Puerto Rico	Trinidad and Tobago	Puerto Rico	Andorra	0.324
	Slovenia	Lithuania	Trinidad and Tobago	Fiji	0.38
	Trinidad and Tobago	Fiji	Suriname	Trinidad and Tobago	0.38
	US Virgin Islands		Albania	Albania	0.408
	Gambia		Moldova	Denmark	0.408
	Cape Verde		Mauritius	Moldova	0.408
			Panama	Mauritius	0.423
			Brunei	Suriname	0.437
			Hong Kong	Lithuania	0.493
			Lesotho	Estonia	0.507
			Eswatini	Croatia	0.521
			Guam	Panama	0.535
			Jamaica	US Virgin Islands	0.535
				Lesotho	0.563
				Jamaica	0.577
				Latvia	0.577
				Brunei	0.592
				Monaco	0.592
Medium Countries					
Zimbabwe	Nicaragua	Belgium	Liberia	Belgium	0
Gabon	Belgium	Hungary	Ecuador	Ecuador	0.097
Guatemala	Sweden	Netherlands	Syria	Sweden	0.113
Nicaragua	Hungary	Sweden	Belgium	Netherlands	0.129
	Netherlands	Liberia	Sweden	Hungary	0.161
	Zimbabwe	Ecuador	Hungary	Liberia	0.161
	Mauritania	Guatemala	Guatemala	Syria	0.29
	Liberia	Syria	Burkina Faso	Nicaragua	0.306
		Burkina Faso	Bulgaria	Bulgaria	0.339
		Bulgaria	Romania	Romania	0.339

Continued

Table 2 Continued

Q1	Q2	Q3	Q4	2020 Scores	
		Romania	Sierra Leone	Burkina Faso	0.387
		Sierra Leone	Nicaragua	Guatemala	0.403
		Switzerland	Zimbabwe	Switzerland	0.452
		Finland	Mauritania	Honduras	0.452
			Greece	Greece	0.468
			Tunisia	Sierra Leone	0.468
			Honduras	Mauritania	0.5
			Haiti	Malawi	0.5
			Malawi	Zimbabwe	0.516
			Togo	Finland	0.532
			Paraguay	Haiti	0.565
			Senegal	El Salvador	0.565
			El Salvador	Tunisia	0.565
			South Sudan		

Large Countries

Q1	Q2	Q3	Q4	2020 Scores	
Philippines	Yemen	Yemen	Yemen	Yemen	0
Sudan	Italy	Italy	Mexico	Italy	0.215
Iran	France	France	Italy	Mexico	0.215
	UK	UK	Sudan	France	0.277
	Spain	Mexico	Iran	UK	0.292
	Algeria	Spain	China	Sudan	0.323
	Mexico	Sudan	Niger	Spain	0.4
	Philippines	Iran	Chad	Iran	0.4
	Sudan	China	Egypt	China	0.431
	Iran	Niger	Bolivia	Chad	0.492
	China	Chad		Egypt	0.508
	Niger	Canada		Algeria	0.523
	Angola	Egypt		Niger	0.523
	Indonesia			Canada	0.538
	USA			Indonesia	0.554
	Chad			Bolivia	0.569
	Canada			Angola	0.585
	Brazil				
	Mali				

Grade F					
Grade D					
Grade C					

states in the USA from 21 December to 31 July, James and Menzies[14] applied HCA method and developed a mathematical framework to compare the trajectories of month-by-month new case counts as a mean of determining effective pandemic responses. They found that there was always three clusters, a big and two smaller clusters of case counts. The stability of the cluster structure over the time gave them a way to look at time-dependent factors and their relation to the states' cluster switching. For example, New York and New Jersey switched from a big to a smaller cluster of the new case counts in May 2020, as they introduced face mask mandates.
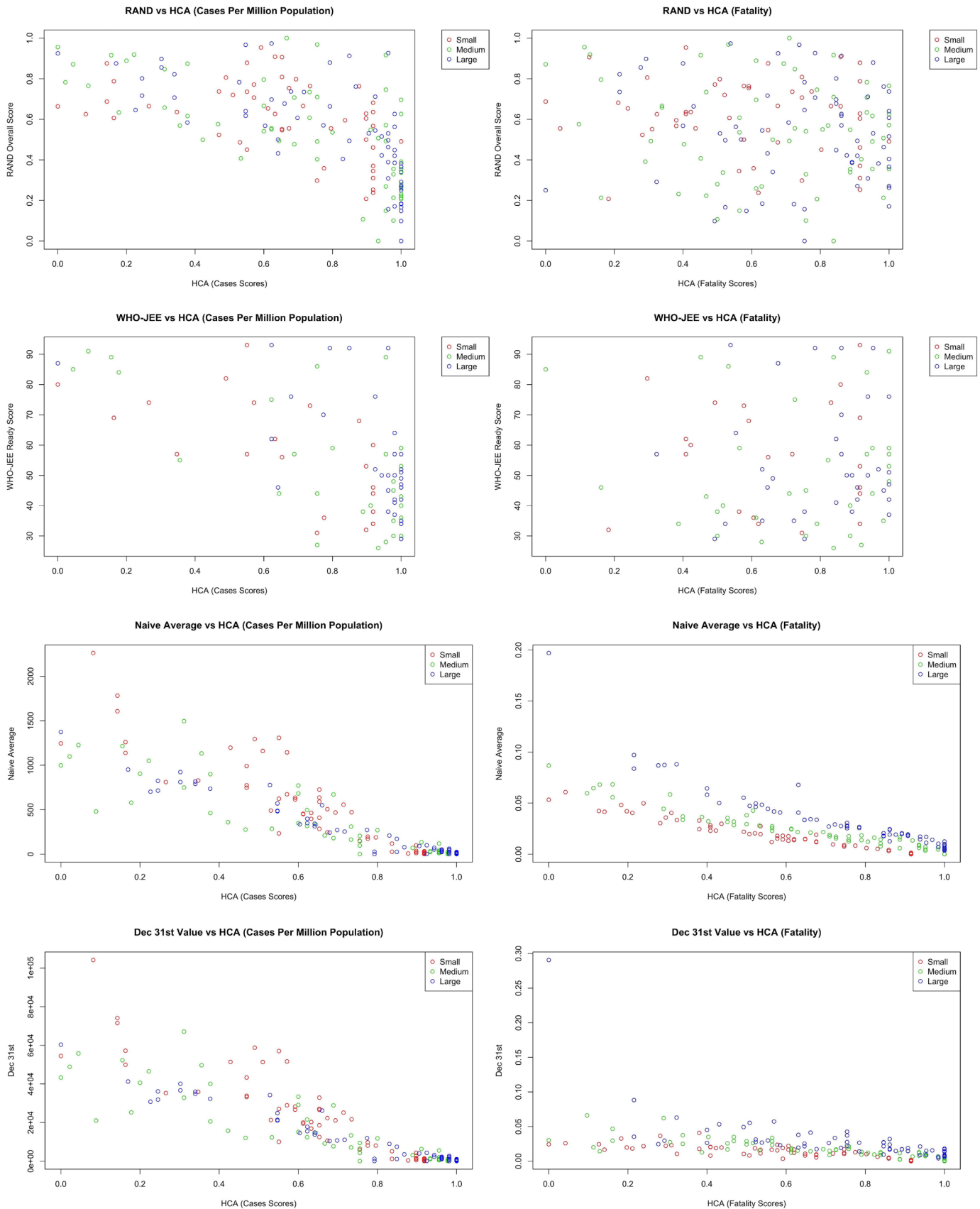
Figure 2 Scatterplot comparisons between RAND, WHO-JEE, naive average, 31 December and hierarchical cluster analysis (HCA).

On the global scale, Zarikas *et al*[15] applied hierarchical clustering on 30 countries, using data from 22 January to 4 April 2020. Mahmoudi *et al*[16] employed fuzzy clustering on seven countries, including France, Germany, Iran, Italy, Spain, the UK and the USA. They focused on daily confirmed cases and deaths and the cumulative cases and deaths. In another related work, Pasin[17] used K-means clustering on total cases and deaths per million, on 191 countries based on the cumulative data as of 22 October 2020. A more interesting approach that uses statistical dimension reduction techniques along with clustering algorithms was proposed by Ramadan *et al*.[18] Their analysis focused on 30 March, 15 April and 25 April data points, including 56, 82 and 91 countries, respectively. They first applied principal component analysis on several COVID-19 related measures, then performed clustering using the Partition around Medoid algorithm. To our knowledge, the only study that explored the emergence of COVID-19 clusters over time, by James and Menzies,[19] used 4 months of daily COVID-19 data for 208 counties. They used a combination of K-means clustering and HCA, with the goal of discovering any critical periods with major policy shifts in the studied countries. However, unlike our method, the number of clusters varied per day, which made it difficult to compare countries' ranking over time.

The results of our retrospective analysis of the COVID-19 pandemic stand in contrast to pandemic preparedness predictions by RAND and WHO-JEE. In RAND's published final report, 22 of the 25 countries most vulnerable to a pandemic are located in sub-Saharan Africa. The other three countries were Haiti, Afghanistan and Yemen. A similar profile of poor access to resources, poor governance and weak health systems was considered as lack of pandemic preparedness.[3] We found no African country being worse than average for COVID-19 cases, but several countries such as Guyana, Gambia and Liberia received the worst grades in COVID-19 fatality. We believe that poor access to resources, poor governance and weak health systems play crucial role in pandemic management resulting in higher fatality, yet the reasons why African countries ranked better than average in COVID-19 cases calls for more research.

The 25 least vulnerable countries in the RAND evaluation had normed scores ranging from 0.82 (Italy) to 1.0 (normed maximum value, for Norway). Overall, 15 of these 25 countries ranked average or worse in our COVID-19 cases analysis. France, Italy and the UK were worse than average in both COVID-19 cases and fatality. We found that countries such as the United Arab Emirates and the USA had excellent outcome in keeping the COVID-19 fatality low; however, their preparedness in controlling the spread of the disease was among the worse in their groups. The interaction between pandemic preparedness and countries' political atmosphere, such as democratic versus

authoritarian governments, is addressed elsewhere[20] and needs to be examined in future research. Meanwhile, we believe that access to the state-of-the-art medical technology, well-trained medical personnel and economic power to provide free of charge COVID-19 patient care could be among some of the important factors in keeping the fatality low in countries like the USA. To explore the complex epidemiological associations between countries' political, social and economic factors, models can benefit enormously by taking advantage of methods developed in computer science and machine learning.

Like RAND and WHO, there are other organisations aiming to develop preparedness measures. For example, The Global Health Security Index,[21] supported by the Bill & Melinda Gates Foundation and the Robertson Foundation, got together 21 experts from 13 countries to create 140 questions, in six categories, and 34 indicators to assess a country's pandemic preparedness. The index uses open-source information available from the world countries. The average of the global health security index among 195 countries is 40.2 out of a possible 100, and the high-income countries score an average of 51.9. In general, the index follows the similar pattern as WHO and RAND results, but it is striking that GHI shows that countries worldwide are not prepared for a globally catastrophic biological event indicating most countries lack foundational health systems capacities vital for epidemic and pandemic response.

Finally in 2019, Oppenheim *et al*[22] published an Epidemic Preparedness Index to assess 188 countries for national-level preparedness. The index consists of five subindices measuring each country's economic resources, public health communications, infrastructure, public health systems and institutional capacity. Similar to the RAND's, their index had a high correlation with the WHO-JEE scores (0.85), and they found that the most prepared countries were concentrated in Europe and North America, while the least prepared countries were clustered in Central and West Africa and Southeast Asia. Their methodology was similar to RAND's approach, using publicly available data such as WHO and World Bank Data, and expert opinion through Delphi's method.

This study suffers from several limitations. The diagnosis and definition of COVID-19 changed over time and varied across countries.[23] For example, some countries included only cases confirmed by PCR tests, while others included presumptive cases when patients displayed clinical symptoms. The fatality may be overestimated or underestimated across the countries depending on the criteria to identify death due to COVID-19 when a patient had other serious comorbidities.[24 25] Although we compared countries within their population and area sizes intentionally, one may argue that we should also adjust for other factors such as countries' economic and sociopolitical situation. We did not explore the impact of those adjustments, as this analysis focuses on unadjusted

COVID-19 outcomes across the world. Perhaps, Yemen with F in COVID-19 fatality and years of war and human tragedy would benefit from such adjustment. Finally, there is no ascertainment of under-reporting or over-reporting of COVID-19 cases or fatality data by countries. Although this may cause a systematic error in the ranking results, the methodology remains robust to apply when more accurate dataset generated in future.

CONCLUSION

This paper is the first to present a COVID-19 performance report for the world through 2020. Our results are in line with those generated by standard epidemiological scoring techniques (naïve average and 31 December values); however, our analysis goes beyond those methods, such as WHO and RAND, to account for countries' behaviour over time and relative to other countries. Existing pandemic preparedness indexes failed to predict the COVID-19 world performance accurately. The complexity of preparedness construct and lack of powerful computational methods may contribute to the shortcoming of existing models, suggesting the need for more work in that area. Future work should also focus on assessing how socioeconomic, political, social and geographic factors, resource availability and policy adoption impacted outcomes.

ORCID iD
Banafsheh Sadeghi http://orcid.org/0000-0003-4611-740X

REFERENCES

1 World Health Oranization. Joint external evaluation tool: international health regulations, 2018. Available: https://apps.who.int/iris/bitstream/handle/10665/259961/9789241550222-eng.pdf?sequence=1 [Accessed 9 Aug 2021].
2 Moore M, Gelfeld B, Okunogbe A, et al. Identifying future disease hot spots: infectious disease vulnerability index. *Rand Health Q* 2017;6:5.
3 Moore M, Gelfeld B. *Adeyemi Theophilus Okunogbe, and Christopher Paul,. identifying future disease hot spots: infectious disease vulnerability index*. Santa Monica, CA: RAND Corporation, 2016.
4 World Health Oranization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020.
5 Worldometer. Coronavirus worldwide graphs. Available: https://www.worldometers.info/coronavirus/worldwide-graphs/#total-deaths [Accessed 9 Aug 2021].
6 Hale T, Angrist N, Goldszmidt R, et al. A global panel database of pandemic policies (Oxford COVID-19 government response tracker). *Nat Human Behav* 2021;5:529–38.
7 Joint External Evaluation (JEE) scores. Available: https://preventepidemics.org/wp-content/uploads/excel/all-countries.xlsx [Accessed 9 Aug 2021].
8 Ward JH. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 1963;58:236–44.
9 Jamison DT, Lau LJ, Wu KB, et al. Country performance against COVID-19: rankings for 35 countries. *BMJ Glob Health* 2020;5:e003047.
10 Kumar S. Use of cluster analysis to monitor novel coronavirus-19 infections in Maharashtra, India. *Indian J Med Sci* 2020;72:44–8.
11 Pooja Sengupta BG, SenRoy S. An analysis of COVID-19 clusters in India-Two case studies on Nizamuddin and Dharavi. *Research Square* 2020.
12 Nascimento MLF. A multivariate analysis on spatiotemporal evolution of Covid-19 in Brazil. *Infect Dis Model* 2020;5:670–80.
13 K-means clustering of COVID-19 cases in Indonesia's Provinces."proceedings of the international conference on global optimization and its applications; 2020 November 21-22.
14 James N, Menzies M. COVID-19 in the United States: trajectories and second surge behavior. *Chaos* 2020;30:091102.
15 Zarikas V, Poulopoulos SG, Gareiou Z, et al. Clustering analysis of countries using the COVID-19 cases dataset. *Data Brief* 2020;31:105787.
16 Mahmoudi MR, Baleanu D, Mansor Z, et al. Fuzzy clustering method to compare the spread rate of Covid-19 in the high risks countries. *Chaos Solitons Fractals* 2020;140:110230.
17 Pasin OaP T. Clustering of countries in terms of deaths and cases of COVID19. *J Health Soc Sci* 2020;5:587–94.
18 Ramadan A, Kamel A, Taha A, et al. A multivariate data analysis approach for investigating daily statistics of countries affected with COVID-19 pandemic. *Heliyon* 2020;6:e05575.
19 James N, Menzies M. Cluster-based dual evolution for multivariate time series: analyzing COVID-19. *Chaos* 2020;30:061108.
20 Kavanagh MM, Singh R, Democracy SR. Democracy, capacity, and coercion in pandemic response: COVID-19 in comparative political perspective. *J Health Polit Policy Law* 2020;45:997–1012.
21 GHS Index. Available: http://www.ghsindex.org [Accessed 9 Aug 2021].
22 Oppenheim B, Gallivan M, Madhav NK, et al. Assessing global preparedness for the next pandemic: development and application of an epidemic preparedness index. *BMJ Glob Health* 2019;4:e001157.
23 Wiersinga WJ, Rhodes A, Cheng AC, et al. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. *JAMA* 2020;324:782–93.
24 Richardson S, Hirsch JS, Narasimhan M, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City Area. *JAMA* 2020;323:2052–9.
25 Weinberger DM, Chen J, Cohen T, et al. Estimation of excess deaths associated with the COVID-19 pandemic in the United States, March to May 2020. *JAMA Intern Med* 2020;180:1336–44.