Citizen Behavior & Health Indicators in Israel During COVID-19: A Systematic Analysis of Data Over Time

Ron S. Kenett[1] *and Carmit Rapaport*[2]

[1]Samuel Neaman Institute, Technion, and KPA, Israel and University of Turin, Italy – ron@kpa-group.com
[2]Institute for Regulation of Emergency and Disaster, College of Law and Business and Department of Geography and Environmental Studies, University of Haifa, Israel – carmit.rapaport@gmail.com

The management of the COVID-19 pandemic involves the interaction and collaboration of three key players: the citizens, policy and decision makers and the healthcare system. The main goal is to find the fragile balance between keeping an open society, a working economy and saving lives. SARS-COV-2 is known to be highly contagious, with a potential for severe health complications, especially in the elderly. In this paper we analyze changes in Israel, in health outcomes and citizen behavior, from the initial stages of the pandemic (March 2020) until the beginning of mass vaccination (December 2021). This perspective includes the impact of government's policies (lockdowns and re-openings) on levels of morbidity. We used official national health data and Google Mobility data to examine these changes. Several models are used to analyze their mutual effects. Results show that public behavior changed over time to a lesser degree than health related data. The implications to policy formulation are also discussed. The main point in the paper is that data driven policy is much needed in conditions of pandemics.

1 Introduction

COVID-19 has led to unprecedented changes in various domains including economic, social, labor, education, air travel, political and technological sectors. Given that the SARS-COV-2, or COVID-19, is highly contagious, a critical goal in this pandemic management has been minimizing the burden on healthcare systems. A related universal goal has been to prevent hospital surge in numbers of critical and ventilated patients.

The data we consider here are preceding the start of vaccination in Israel. The early indications are that vaccinations are effective and match the vaccination efficiency estimates from the clinical trials. However, many uncertainties remain regarding resilience to mutations and contagion effects of vaccinated populations.

Pandemic management policies in Israel included national and local lockdowns, with frequent closures of air traffic, education institutions, and economic

DOI: 10.54103/milanoup.73.51

sectors. Citizens were called to adhere to protection measures, such as keeping physical distance, practicing hygiene measures and using face masks. Government applied policies such as lockdowns and re-openings according to the perceived “acceptable loss" derived from a balance between running the economy and minimizing economic and social damage versus the need to save lives.[1] As expected, easing of lockdowns led to an increase in morbidity. At the same time, lockdowns and quarantines caused severe damage in terms of social and economic factors as reflected by increasing rates of domestic violence, unemployment, depression, non-normative behavior (alcohol drinking and drugs) among others.[2–4]

Research has shown that citizen compliance with stay-at-home policies[5] is predicted by perceived risks and trust in science and scientists,[4] trust in the authorities[6] and social capital.[7] Moreover, restricted mobility options (such as closures of air traffic, shopping malls, sport events and education institutions), were found to effectively prevent outbreaks and reduce the number of death.[8–10] Furthermore, research has also found that citizens voluntarily decreased their mobility, without official instructions to stay at home.[11,12]

In the current analysis, we examine both health and behavior, or activity related data of Israeli citizens, over time. Israel experienced national lockdowns in March, September and December 2020. Referring to official national health data and to Google Mobility data we trace patterns in citizens behavior, in response to lockdowns and re-opening, as well as changes in different behavior domains such as attendance at workplace, public places and staying at home behavior, in response to changing reported health indicators and government policies.

2 Methodology

We used Israeli official Ministry of Health daily data from March 1st, 2020 until Dec. 28th, 2020. Variables included in the study are: number of tests, number of hospitalized patients, number of people ventilated, number of reported positive cases, number of severe condition patients, and number of COVID19 attributed death, per day. We also use Google Mobility data (https://www.google.com/covid19/mobility/) for Israel for the same time frame. Variables included mobility indicators for retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, and residential places. Mobility variables show the percent change from pre-pandemic baseline.

The analysis included multivariate T^2 control charts, cluster analysis, decision trees and structural equation models (SEM). To account for the nonstationary in the data, ARIMA models were used first and the multivariate control charts were applied to the residuals. The analysis was performed using the JMP 15.2 software (http://www.jmp.comwww.jmp.com).

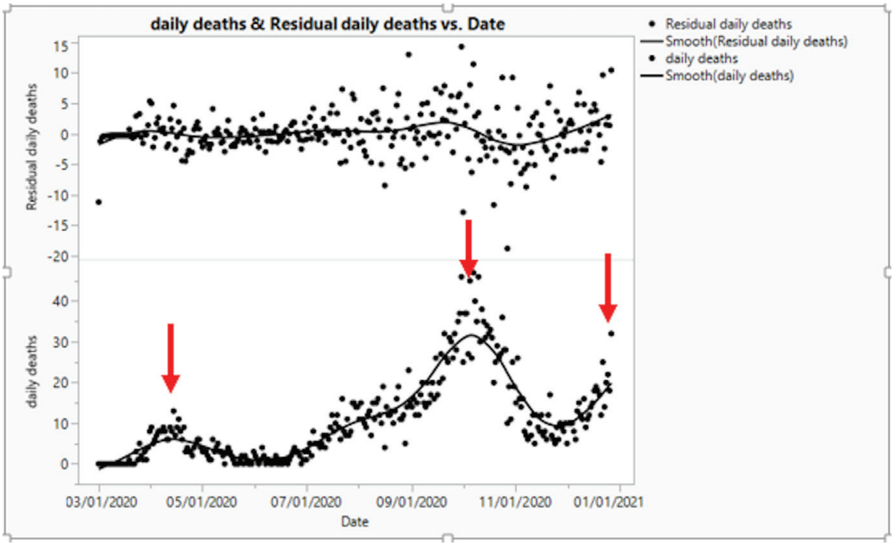


Figure 1: Number of daily deaths and residuals from AR(8), over time.

3 Findings

3.1 Health data during COVID-19

The lower panel in Figure 1 is the number of daily deaths in Israel, over time. The first, second and third pandemic waves are clearly visible when adding a smoother to the raw data. The data is non stationary and, in order to apply control charts which assume independence between observations, we fit an autoregressive model of type AR(8) to all the health and activity data. The AR(8) model was found to best fit the 12 time series. The upper panel in Figure 1 shows the residuals of daily deaths, after fitting the AR(8). Some heteroscedasticity is apparent towards the end of the series but this does not affect the analysis we conducted with non parametric methods such as K-means clustering and decision trees. In order to handle the combined 6 health related indicators we apply a multivariate T^2 control chart to the 6 series of AR(8) residuals.[13] In that chart, the Mahalanobis T^2 distance is used to indicate deviations from the 6-dimensional vector of averages of residuals over the 303 observations. The distance is calibrated to account for the correlation structure of these variables. The T^2 control chart on these 6 dimensions is presented in Figure 2 .

In order to identify clusters in the health data patterns, a K-means clustering of the health data was conducted with K=4. This was an optimal clustering using the cubic clustering criterion (CCC) as shown in Figure 3. The CCC is used to estimate the number of clusters using Ward's minimum variance method, K-means, or other methods based on minimizing the within-cluster sum of squares.

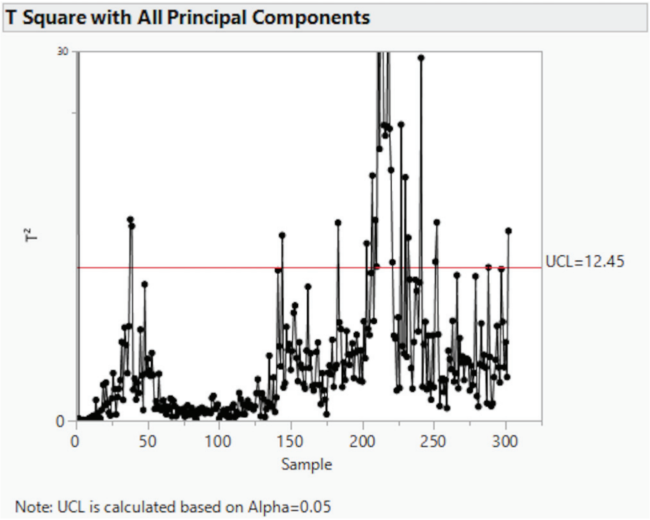


Figure 2: T^2 chart on six health related variables.

Iterative Clustering

Cluster Comparison

Method	NCluster	CCC	Best
K Means Cluster	4	5.31627	Optimal CCC

Columns Scaled Individually

K Means NCluster=4

Columns Scaled Individually

Cluster Summary

Cluster	Count	Step	Criterion
1	111	11	0
2	86		
3	61		
4	44		

Cluster Means

Cluster	daily tests	hospitalized	ventilated	confirmed COVID19	severe condition	daily deaths
1	9738.38739	246.63964	36.0630631	14871.4144	54.0630631	1.92792793
2	21560.8372	858.581395	106.860465	71912.8605	309.767442	11.0348837
3	54596.8033	783.114754	119.95082	335434.885	366.819672	12.2459016
4	41128.25	1503.81818	191.022727	263568.227	691.045455	31.0681818

Figure 3: K-means analysis of six health related variables, with K= 4.

The performance of the CCC is evaluated by Monte Carlo methods. The results show that clusters vary in the volume of the health components in each cluster and its type, see Figure 3. For example, Cluster 1 includes lower levels in all health indicators in comparison with the other clusters. Cluster 3, represents

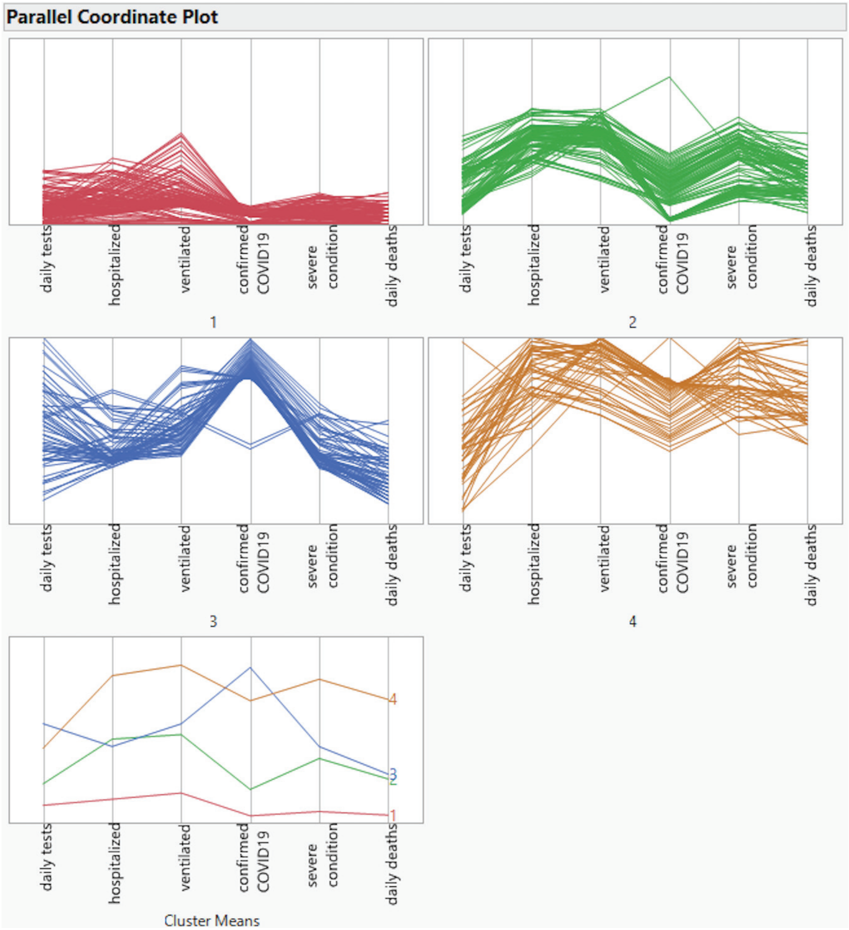


Figure 4: Parallel plots of the 4 clusters of the health-related data.

extreme levels of daily tests and confirmed (positive) cases, with non severe data in other health indicators, as in Cluster 4.

A parallel plot display of the 4 clusters is shown in Figure 4. The increase in confirmed cases is characteristic of Cluster 3. This was low in Clusters 1 and 2. Higher levels of morbidity were found in Cluster 4. These results indicate different patterns of the pandemic over time.

The derived 4 clusters are consistent with time ordering, see Figure 5. Cluster 1 is the initial phase with cluster 2 indicating the first pandemic wave. Cluster 4 is the second wave and Cluster 3 is the third wave.

A complementary analysis was conducted by replacing the health raw data with three ratios representing the effect of exposure. The indicators we calculated are *'pandemic outbreak'* measured as the number of positive cases divided by number

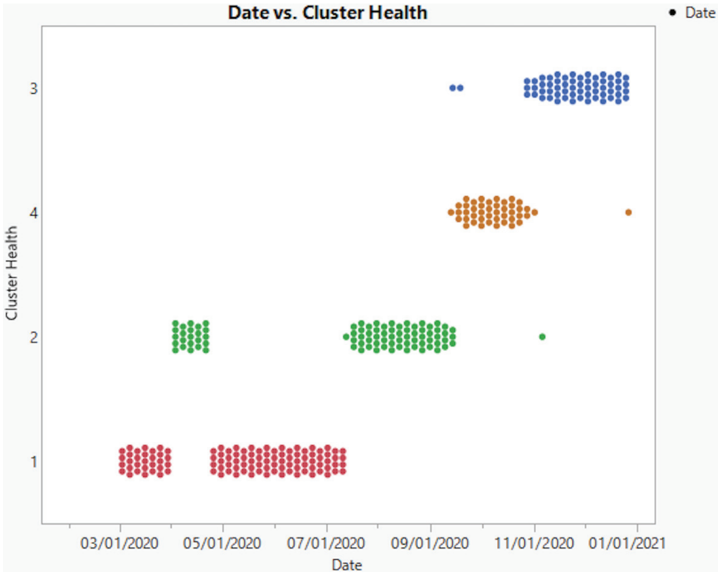


Figure 5: The 4 clusters of the health-related data (note reordering of clusters 3 and 4 to match the time dimension).

of testes, *'healthcare system's capacity'* measured by number of severe conditioned patients divided by the number of hospitalized, and *'pandemic severity'* which is measured by the number of deaths per number of patients hospitalized. Figure 6 present these three indicators over time.

From Figure 6 we see in Cluster 3 a stable ratio of confirmed/tested with an increased level of severe/hospitalized and deaths/hospitalized in Clusters 4 and 3 relative to Clusters 1 and 2.

Control charts of severe/hospitalized and deaths/hospitalized are presented in Figure 7. They show that the patients with severe conditions, as a proportion of hospitalized patients, has increased significantly from Clusters 1 and 2 to Clusters 2 and 4 and that the proportion of COVID related deaths as a proportion of hospitalized patients is reduced in Cluster 3 relative to Cluster 4. For an application of control charts to survey data analysis, see.[14] For general applications see.[13]

The results shed light on the changes in the pandemic over time. However, these changes vary across indicators. In other words, the pandemic does show increases in all parameters at once, but rather, is reflected in the combination of several indicators.

3.2 Public behavior and activity during COVID-19

We proceed to analyze the activity related data. Figure 6 is the T^2 chart for the activity data residuals after fitting an AR(8). Comparing Figure 8 to Figure 2

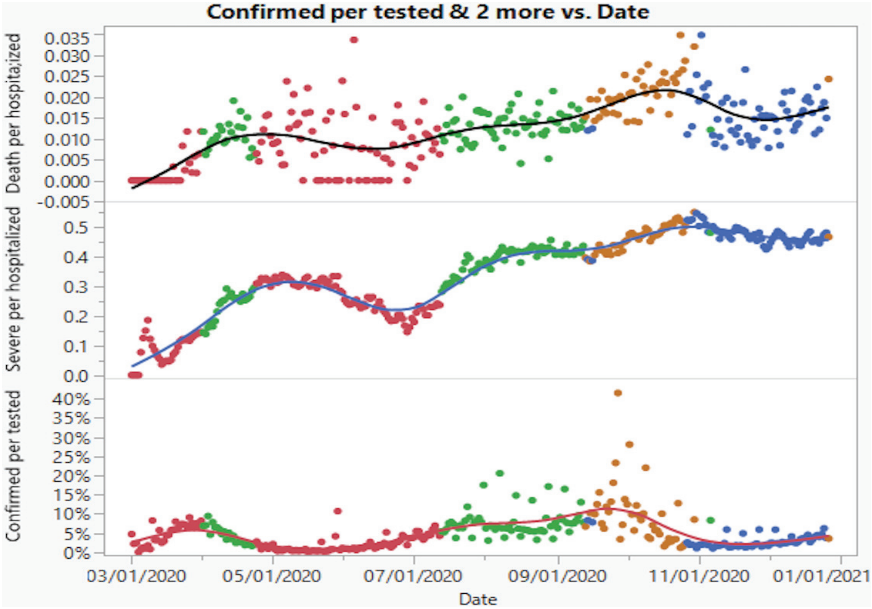


Figure 6: Trends of confirmed/tested, severe/hospitalized and deaths/hospitalized.

we observe different patterns. Clusters 2 and 4 corresponding to the second and third pandemic waves, did not show parallel behavior in the activity chart. In other words, the behavior of citizens in the period of clusters 2 and 4 apparently ignored the trends in health-related data. Moreover, government policy - of national stay-at-home orders in this time period proved less effective.

A classification predictive analysis, with a random forest analysis of the activity related data as predictor of the health clusters, produces a 19% misclassification rate in the confusion matrix. The ROC curves are shown in Figure 9 with AUC ranging from 93% to 97%. The “column contributions" shows that the clusters are mostly determined by the decrease in transportation and retail activity. Less so by the other 4 activity related variables (Figure 10). Closure of public transportation and retail activity are included in the government’s policy, and therefore forced the public to avoid it. However, grocery and pharmacy and visiting parks were less effected - as the decrease in these activities were minor.

As the decision tree in Figure 11 indicates, the decrease in transit stations is the significant reaction of all health-indicators clusters. Furthermore, as the left leaf in the figure shows, for Cluster 1, (the first pandemic-wave characterized by lower levels of morbidity, but the beginning of the pandemic), the decrease was also in retail, then grocery and pharmacy and workplaces. However, in the 2nd and 4th cluster, which represent the second and third waves of the pandemic when morbidity was very high, we see a different behavioural pattern: transportation

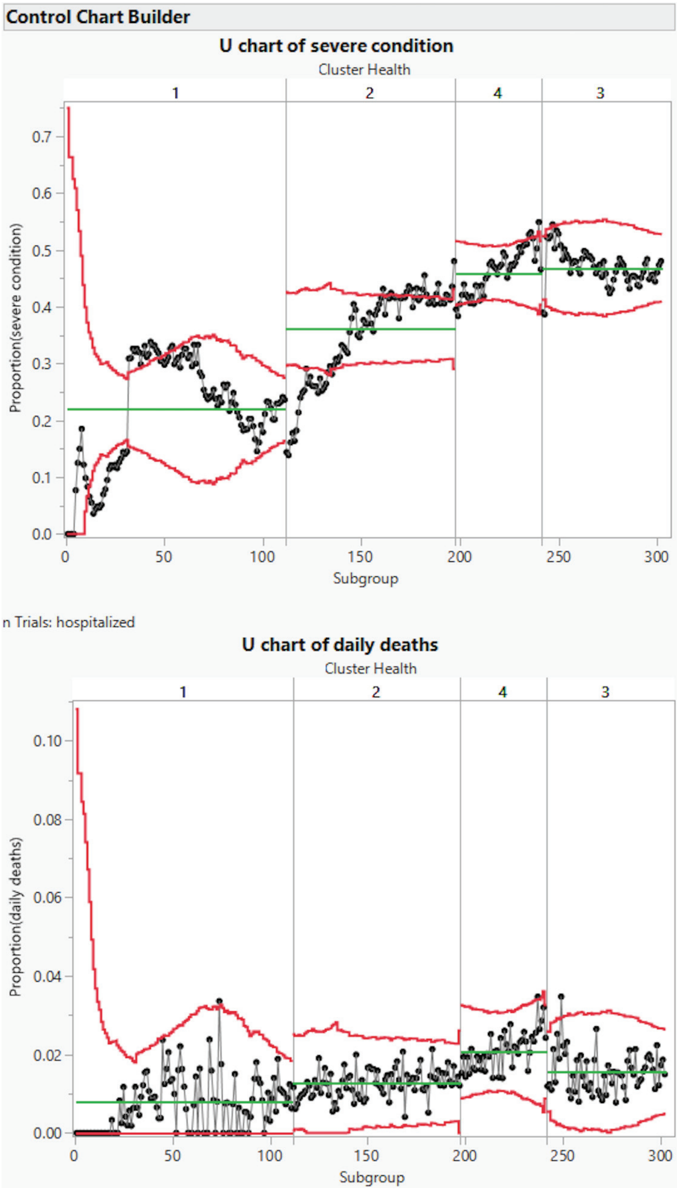


Figure 7: U charts of severe/hospitalized and deaths/hospitalized by cluster.

and retail are decreased by the force of the national lockdown, however the grocery and pharmacies are increasing as well as workplace attendance. This means that although the morbidity was high, with increasing numbers of cases of deaths people hospitalized in severe condition, the public's behaviour was not compliant with lockdown orders.

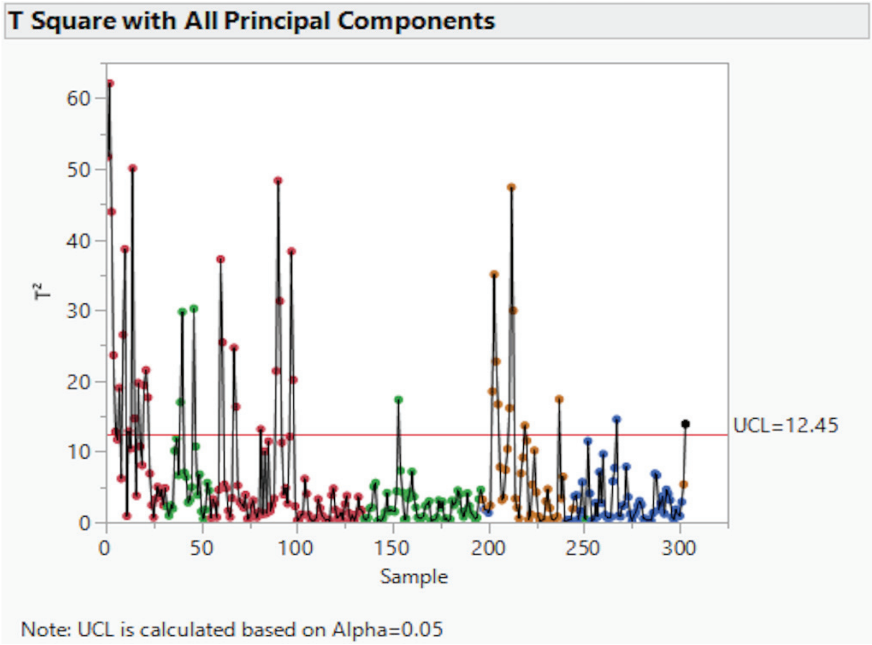


Figure 8: T^2 chart on six activity related variables.

The tree with 10 splits is presented in Figure 11 followed by a listing of its leaves Figure 12

Yet another analysis aiming at identifying latent variables affecting the health and activity data was conducted using confirmatory structural equation models (SEM) applied to the standardized health residuals and the activity residuals, see Figure 13.

The SEM comparative fit index (CFI) was marginal (0.79) and only the loadings on the citizen activity latent variables turned out significant with a nonsignificant covariance of the health and behavior latent variables of -0.017. CFI represents the model fit by examining the discrepancy between the data and the hypothesized model while adjusting for sample size inherent in the chi-squared test of model fit. CFI values range from 0 to 1, with larger values indicating better fit. This reflects the findings in Figure 5, 6 and 7 which indicated that the health variable change over time, but separately - not all health indicators increased at the same time. Furthermore, we run another analysis where we added to the model regressions between the health indicators and each of the behavior variables. Results showed an improved fit (CFI=0.91).

The loadings (Figure 14) show that increased number of hospitalized patients increases visits to grocery stores and pharmacy and decreases attendance at workplaces. Also, confirmed cases had the strongest effect on decreased attendance at

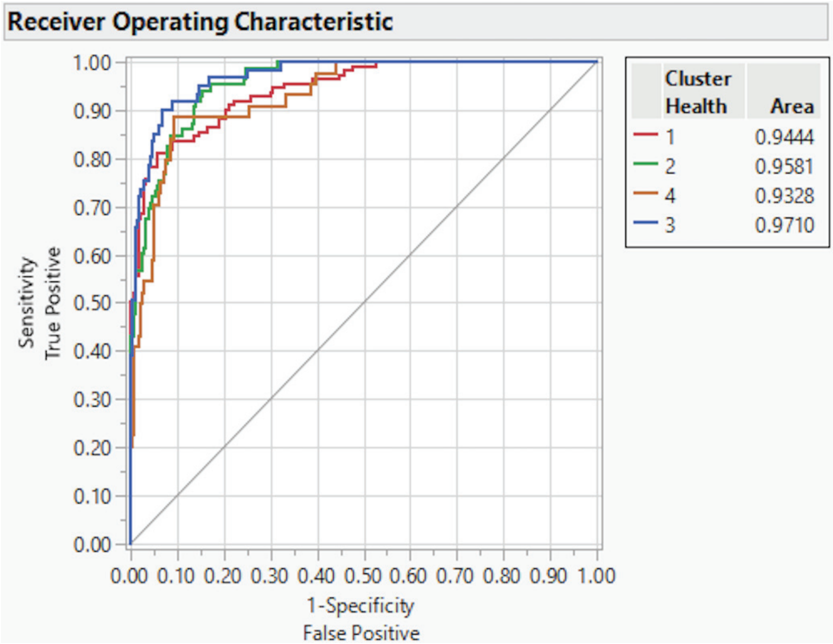
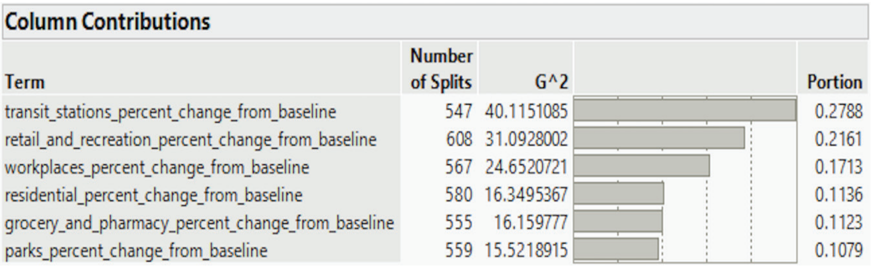


Figure 9: ROC curves from random forest predictive model of the activity data as predictors of the health related clusters.



Column Contributions

Term	Number of Splits	G^2		Portion
transit_stations_percent_change_from_baseline	547	40.1151085		0.2788
retail_and_recreation_percent_change_from_baseline	608	31.0928002		0.2161
workplaces_percent_change_from_baseline	567	24.6520721		0.1713
residential_percent_change_from_baseline	580	16.3495367		0.1136
grocery_and_pharmacy_percent_change_from_baseline	555	16.159777		0.1123
parks_percent_change_from_baseline	559	15.5218915		0.1079

Figure 10: Column contributions in random forest predictive model of the activity data as predictors of the health related clusters.

workplaces. The number of daily deaths did not affect the behavior. The findings show that the number of hospitalized and ventilated patients were predictors of behavioral change.

4 Discussion

The paper provides an analysis of daily health data and citizens' activity related data (March - December 2020) in Israel during the COVID-19 pandemic.

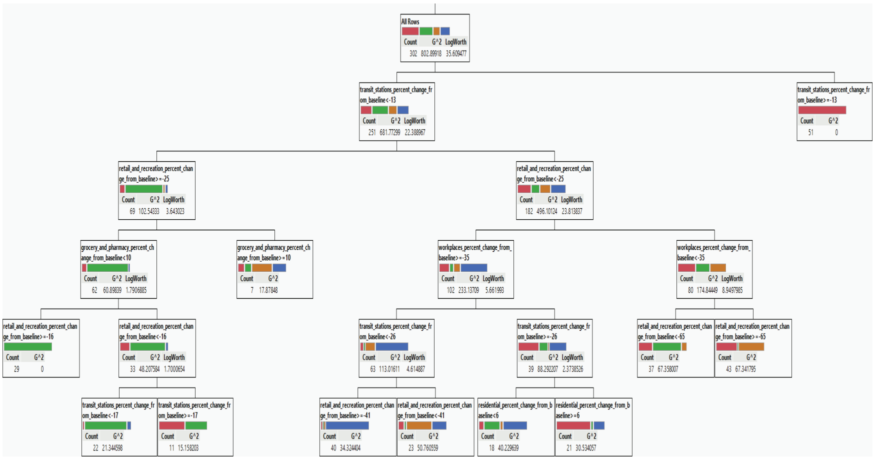


Figure 11: Decision tree predictive model with 10 splits of the activity data.

Multivariate control charts were used to monitor changes over time and three pandemic waves are visually apparent. A K-means clustering analysis identifies 4 health-indicators clusters that are consistent with the three pandemic waves over time. The clusters are characterized by lower levels of morbidity at the beginning of the pandemic (March 2020), then moderate levels of morbidity with relatively low number of tests and confirmed cases (September 2020). Two clusters (3+4) representing higher levels of morbidity with relatively low numbers of tests and positive cases (Cluster 3). This indicates a variability in the term "morbidity". This is an important aspect for pandemic management and policy. The question posed, in that context, is not how high are the different indicators, but rather, who are the important indicators, or combination of indicators, which represent an extreme situation that demands extreme actions.

The changes in health-related indicators over time and the cluster which were found in this study, are followed by changes in public activity, such as adherence to the stay at home instructions. Governmental policies such as national lockdowns were very effective in the first wave. The results show lower morbidity rates and significant decrease in public activity, with an increase in grocery and pharmacy visits. However, over time, in the second and third waves of the pandemic, the same policies did not achieve the expected results. The public did not change its behavior in accordance with the high morbidity and vice versa - the increased activity increased the morbidity rates. The change in public activity which were officially supposed to be locked, such as retail and recreation, was less significant in the second and third waves than in the first wave. This represents an adaptive behavior pattern of the public to the unfolding situation. A closer examination of the results found that as the health indicators varied across the pandemic

Leaf label	Cluster 1 (red)	Cluster 2 (green)	Cluster 4 (brown)	Cluster 3 (blue)
Transit stations<-13, retail and recreation>=-25, grocery and pharmacy<10, retail and recreation>=-16	0.0103	0.9797	0.0042	0.0059
Transit stations<-13, retail and recreation>=-25, grocery and pharmacy<10, retail and recreation<-16, transit stations<-17	0.0565	0.8445	0.0049	0.0941
Transit stations<-13, retail and recreation>=-25, grocery and pharmacy<10, retail and recreation<-16, transit stations>=-17	0.5250	0.4519	0.0094	0.0137
Transit stations<-13, retail and recreation>=-25, grocery and pharmacy>=10	0.1664	0.1674	0.3923	0.2739
Transit stations<-13, retail and recreation<-25, workplaces>=-35, transit stations<-26, retail and recreation>=-41	0.0563	0.0302	0.0283	0.8852
Transit stations<-13, retail and recreation<-25, workplaces>=-35, transit stations<-26, retail and recreation<-41	0.1379	0.0516	0.5066	0.3039
Transit stations<-13, retail and recreation<-25, workplaces>=-35, transit stations>=-26, residential<6	0.1234	0.3291	0.0601	0.4874
Transit stations<-13, retail and recreation<-25, workplaces>=-35, transit stations=-26, residential>=6	0.6975	0.0569	0.0064	0.2391
Transit stations<-13, retail and recreation<-25, workplaces<-35, retail and recreation<-65	0.2987	0.5863	0.1098	0.0051
Transit stations<-13, retail and recreation<-25, workplaces<-35, retail and recreation>=-65	0.4398	0.0291	0.5267	0.0044
Transit stations>=-13	0.9878	0.0055	0.0028	0.0039

Figure 12: Leaves of the decision tree of the activity data.

waves, the public tended to decrease its activity as a reaction to increasing levels of positive cases and the number of hospitalized. This represents the way the public perceives the severity of the pandemic and how it acts accordingly.

Lockdowns and stay-at-home instructions in Israel had an impact on citizen activity. However, at the beginning of the events (March 2020), public compliance was much more significant than in the next two waves and lockdowns. These results provide policy makers a method and insights for tracking public compliance with stay-at-home instructions, when facing various levels of morbidity.

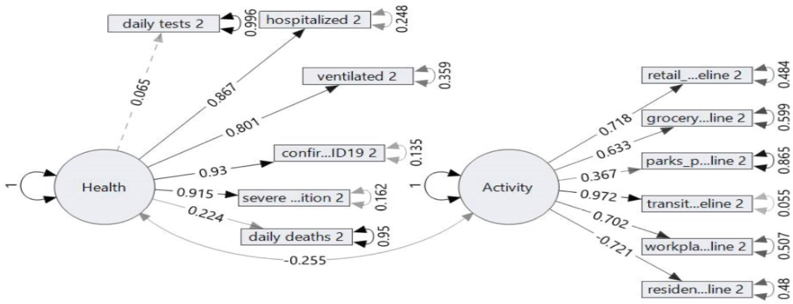


Figure 13: Structural equation model analysis of the health and activity variables with latent variables behind the health and activity data.

Regressions	Estimate	Std error	Wald Z	Prob >Z
daily tests 2 → retail and recreation	0.100314	0.056932	1.761994	0.0781
daily tests 2 → grocery and pharmacy	0.040184	0.058529	0.686561	0.4924
daily tests 2 → parks	-0.04888	0.059814	-0.81726	0.4138
daily tests 2 → transit stations	0.157795	0.057094	2.763801	0.0057
daily tests 2 → workplaces	0.108495	0.05643	1.922643	0.0545
daily tests 2 → residential	-0.12688	0.058122	-2.18301	0.029
hospitalized 2 → retail and recreation	0.355041	0.207605	1.710177	0.0872
hospitalized 2 → grocery and pharmacy	0.409025	0.200172	2.043363	0.041
hospitalized 2 → parks	0.063684	0.153197	0.4157	0.6776
hospitalized 2 → transit stations	0.231329	0.207474	1.114976	0.2649
hospitalized 2 → workplaces	-0.59716	0.216911	-2.753	0.0059
hospitalized 2 → residential	-0.03614	0.165549	-0.2183	0.8272
ventilated 2 → retail and recreation	0.051492	0.142212	0.362082	0.7173
ventilated 2 → grocery and pharmacy	-0.01365	0.13915	-0.0981	0.9219
ventilated 2 → parks	0.092759	0.114601	0.809403	0.4183
ventilated 2 → transit stations	-0.04009	0.142314	-0.2817	0.7782
ventilated 2 → workplaces	-0.46903	0.147282	-3.18456	0.0014
ventilated 2 → residential	0.023281	0.120022	0.193976	0.8462
confirmed COVID19 2 → retail and recreation	-0.4175	0.359958	-1.15986	0.2461
confirmed COVID19 2 → grocery and pharmacy	-0.12312	0.342118	-0.35988	0.7189
confirmed COVID19 2 → parks	-0.18703	0.239873	-0.77971	0.4356
confirmed COVID19 2 → transit stations	-0.22	0.359224	-0.61244	0.5402
confirmed COVID19 2 → workplaces	-1.33665	0.387411	-3.45022	0.0006
confirmed COVID19 2 → residential	0.164108	0.269891	0.608053	0.5432
severe condition 2 → retail and recreation	-0.1152	0.308709	-0.37318	0.709
severe condition 2 → grocery and pharmacy	-0.1654	0.294237	-0.56214	0.574
severe condition 2 → parks	0.032263	0.21	0.153631	0.8779
severe condition 2 → transit stations	-0.08664	0.308175	-0.28114	0.7786
severe condition 2 → workplaces	-1.08946	0.330702	-3.29438	0.001
severe condition 2 → residential	-0.13512	0.234305	-0.57667	0.5642
daily deaths 2 → retail and recreation	-0.09566	0.057844	-1.65369	0.0982
daily deaths 2 → grocery and pharmacy	-0.08587	0.059343	-1.44698	0.1479
daily deaths 2 → parks	-0.10376	0.060163	-1.7247	0.0846
daily deaths 2 → transit stations	-0.07211	0.058004	-1.24313	0.2138
daily deaths 2 → workplaces	-0.10003	0.056088	-1.78347	0.0745
daily deaths 2 → residential	0.096292	0.058589	1.643518	0.1003

Figure 14: Regressions of SEM model.

References

[1] Ashkenazi, I., Rapaport, C. (2020). Saving Lives Versus Saving Dollars: The Acceptable Loss for Coronavirus Disease 2019. *Critical Care Medicine*, 48(8): 1243–1244.

[2] Bradbury-Jones, C., Isham, L. (2020). The pandemic paradox: the consequences of COVID-19 on domestic violence. *Journal of Clinical Nursing*, 29(14–14): 2047–2049.

[3] Kawohl, W., Nordt, C. (2020). COVID-19, unemployment, and suicide. *The Lancet Psychiatry*, 7(5): 389–390.

[4] Plohl, N., Musil, B. (2021). Modeling compliance with COVID-19 prevention guidelines: The critical role of trust in science. *Psychology, Health and Medicine*, 26(1): 1–12.

[5] Engle, S., Stromme, J., Zhou, A. (2020). Staying at home: mobility effects of covid-19. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3565703.

[6] Bargain, O., Aminjonov, U. (2020). Trust and compliance to public health policies in times of COVID-19. *Journal of Public Economics*, 192: 104316.

[7] Borgonovi, F., Andrieu, E. (2020). Bowling together by bowling alone: Social capital and Covid-19. *Social Science and Medicine*, 265: 113501.

[8] Kraemer, M. U., Yang, C. H., Gutierrez, B., Wu, C. H., Klein, B., Pigott, D., Brownstein, J. S. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368(6490): 493—497.

[9] Yilmazkuday, H. (2020). COVID-19 spread and inter-county travel: Daily evidence from the US. *Transportation Research Interdisciplinary Perspectives*, 8: 100244.

[10] Yilmazkuday, H. (2020). Stay-at-home works to fight against COVID-19: international evidence from Google mobility data. *Journal of Human Behavior in the Social Environment*, 31(1-4): 210–220.

[11] Barrios, J. M., Benmelech, E., Hochberg, Y. V., Sapienza, P., Zingales, L. (2021). Civic capital and social distancing during the Covid-19 pandemic. *Journal of Public Economics*, 193: 104310.

[12] Maloney, W. F., Taskin, T. (2020). Determinants of Social Distancing and Economic Activity During COVID-19: A Global View. World Bank Policy Research Working Paper, (9242).

[13] Kenett, R. S., Zacks, S. (2021). *Modern Industrial Statistics: With Applications in R, Minitab and JMP (3rd ed.)*. New York, NY: John Wiley and Sons.

[14] Kenett, R.S., Salini, S. (2011). *Modern Analysis of Customer Surveys: with applications using R*. New York, NY: John Wiley and Sons.