Integrated Analysis of Behavioral and Health Data: A Comparative Study of COVID19 Data in Israel and Italy

Ron Kenett (✉ ron@kpa-group.com)

Samuel Neaman Institute for National Policy Research

Giancarlo Manzi

University of Milan

Carmit Rapaport

University of Haifa

Silvia Salini

University of Milan

Research Article

Keywords: COVID19, mitigation policy responses, Italy, Israel

Posted Date: October 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-892584/v2>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Integrated Analysis of Behavioral and Health Data: A Comparative Study of COVID19 Data in Israel and Italy

Ron Kenett^{1,*}, Giancarlo Manzi², Carmit Rapaport³ and Silvia Salini²

¹KPA Group and Samuel Neaman Institute, Technion, P.O. Box 2525, Raanana, 43100, Israel

²Department of Economics, Management and Quantitative Methods and Data Science Research Center, University of Milan, Milan, 20122, Italy

³Department of Geography and Environmental Studies, University of Haifa, Haifa, 3498838, Israel and NIREG-Institute for Regulation of Emergency and Disaster, College of Law and Business, Ramat Gan, 5110801, Israel

*ron@kpa-group.com

ABSTRACT

The response to the COVID19 pandemic has been highly variable, both in terms of between-nations variation and within the same nation, at different waves. In this context, governments applied different mitigation policy responses with varying impact on social and economic measures over time. This article examines the effect of mobility restriction measures in Italy and Israel and compares the association between health and population mobility data. Facing the pandemic, Israel and Italy implemented different policy measures and experienced different public activity patterns. The analysis we conducted is a staged approach using Bayesian Networks and Structural Equations Models to investigate these patterns. The goal is to assess the impact of pandemic management and mitigation policies on pandemic spread and population activity. We propose a methodology that first models data from health registries and Google mobility data and then shows how decision makers can conduct scenario analysis to help support pandemic management policies.

Introduction

The COVID19 pandemic has far-reaching consequences on global, national and local economies. Management of the pandemic is affected by national policies. These policies include movement restrictions and massive testing to detect outbreaks. However, given the human-to-human transmission of the SARS-CoV-2, an important factor in preventing outbreaks is the population's behavior and compliance with official instructions. Public's adherence to the instructions, such as wearing face masks, keeping physical distance and hygiene practice compliance, requires a behavioral change which cannot be taken for granted. Strictly enforcing it at a national level, is almost impossible. Furthermore, disruptive pandemic management policies, such as national and local lockdowns, with frequent closures of air traffic, education institutions, and economic sectors have significant economic and social consequences. These include increasing rates of domestic violence (Bradbury-Jones & Isham, 2020), unemployment (Kawohl, & Nordt, 2020), depression (Pfefferbaum & North, 2020), psychological distress (Xue et al., 2021), and non-normative or addictive behaviors (alcohol drinking and drugs use among others) (Bonny-Noach & Gold, 2020; Calina et al., 2021). Governments policies such as lockdowns and re-openings are determined according to a perceived "acceptable loss", that is minimizing economic and social damage while also saving lives (Ashkenazi & Rapaport, 2020). It has become clear that pandemic management must take into consideration social, public health, behavioral and economic factors, as well as a constant monitoring of the population's ability to maintain restrictions over time.

Policy and non-pharmaceutical interventions, such as stay at home instructions and closures of businesses and air traffic, have positive effects on morbidity and death rates (Flaxman et al., 2020; Kraemer et al., 2020; Yilmazkuday, 2020), as well as using face masks (Chernozhukov et al., 2021). However, citizens also voluntarily decreased their mobility, without official instructions to stay at home (Barrios et al., 2021; Maloney & Taskin, 2020; Yan et al., 2021), and some groups even objected to the removal of restrictions (Czeisler et al., 2020). Citizens' compliance with stay-at-home policies (Engle et al. 2020) was associated to perceived risks (Dryhurst et al., 2020), trust in science and scientists (Plohl & Musil, 2021), trust in the authorities (Bargain and Aminjonov, 2020), social capital (Borgonovi & Andrieu, 2020) and political orientation (Painter & Qiu, 2021). Also, social responsibility in caring for relatives and community (Wang et al., 2021), future economic status (Bodas & Peleg, 2020) and messages of prosocial advantages of adhering to the instructions (Joradan, Yoeli & Rand, 2020), contributed to higher levels of adherence to lockdown instructions. Taken together, these findings suggest that a range of factors, both at the

individual and the social level, affect citizens' compliance with protection instructions. This emphasizes the importance of integrating behavioral and policy measures in pandemic management.

In this paper we focus on the analysis of data on population behavior and health data such as the number of deaths and patients in intensive care units (ICU). Specifically, we consider the impact of lockdowns and mobility restrictions as reflected by public behavior. The challenge is to combine these sources of data into a unified system that provides support and information to decision makers. The data analyzed is from Israel and Italy where we can combine first-hand knowledge and personal experience. Our goal is to identify links between population behavior (activity), COVID19 health data and policy decisions. The analysis of data from both countries requires calibration since a simple temporal consideration is not adequate. Moreover, the dynamics of the pandemic in the two countries was not synchronized and the policy decisions were different, at different stages of the COVID19 pandemic. In order to achieve such a calibration, we considered waves and lockdowns as comparable conditions and examined overall associations between measured variables in both countries.

We combine a multivariate exploratory analysis, with Bayesian networks and a confirmatory analysis applying structural equation models. Finally, in order to provide the capability to evaluate alternative scenarios, we update the Bayesian networks, following discretization of the original data using specific thresholds applied in Israel and Italy. The thresholds on health indicators were set to safeguard hospitalization capabilities in the two countries.

Pandemic management policies in Italy and Israel

The first case of COVID19 in Israel was discovered on February 28st, 2020. March 10-11, 2020 was the Jewish holiday, Purim, which involves mass gatherings and public celebrations. This caused a sharp increase in the number of the positive cases, which led to drastic movement restrictions in the entire country. On March 14th, 2020 the academic and school systems were closed and gatherings of more than 10 persons were forbidden. On March 25th, 2020 Israel underwent a general lockdown. In April 19th, 2020, due to a decrease in the number of ICU patients and those hospitalized in severe condition, a first ease of the lockdown allowed for sport activities to be practiced by more than 500 meters from one's home. In May 2020, schools, public transportation and public places like restaurants reopened. In July and August, health measures indicated a new outbreak, and new restrictions were set until a second general lockdown which started on September 18th, 2020. On October 17th, 2020, public places and schools started to reopen against recommendations of health professionals. A rising number of daily positive cases, number of deaths and severe patients led to a third wave (starting at the beginning of December 2020). On January 8th, 2021 Israel was under a third national lockdown that lasted one month. Since December 20th, 2020 a massive vaccination operation was conducted with world breaking records in terms of number of vaccinated people. The data considered here is pre-vaccination.

The first COVID19 cases reported in Italy - two Chinese tourists traveling across Italy - were discovered on January 29th, 2020. These cases did not trigger any outbreak as the infected persons were immediately isolated and hospitalized. The first real outbreak was reported in the village of Codogno in Northern Italy. This was the first serious hot-spot of the COVID19 pandemic outside China which spread to Europe and other Western countries, included the US. On March 4th, 2020 all the schools and universities were closed in Italy and on March 11th, 2020 a national lockdown was introduced. On June 3rd, 2020 the national lockdown was eased and free movement between regions was allowed and schools and shops reopened.

Following the summer, when reopening of discos, gyms and free movements for touristic reasons was allowed, in September 2020 a second wave of the pandemic started, and new restrictions were introduced in October 2020. On November 6th, 2020 a regional "traffic light system" was set in Italy according to the current pandemic severity level: (i) "yellow" areas where all shopping malls and supermarkets are closed during weekends. Cultural and recreation places (museums, and cinemas) are closed, and restaurants are open until 6PM. Only elementary and primary schools are open, while universities are open with 50% of the students allowed to attend classes also, a curfew from 10PM to 6PM is set; (ii) "orange" areas where all restaurants are closed all day, public transport is open on a 50% filling capacity level, universities are closed, only elementary and primary schools are open, malls are closed during the weekends, gyms, cinemas theatres and museums are closed; curfew is still in place and movements are allowed only locally; (iii) "red" areas where no movements are allowed apart from emergency and specialcases, all shops, school, universities, restaurants, cinemas, theatres, museums, gyms are closed, with the exception of essential services; curfew from 10PM to 6AM is still in place. In mid-December 2020 the third wave of pandemic started, leading in March 2021 to a situation where almost all the Italian regions were in the "red area" alert.

On December 27th, 2020 the first vaccination jab was given in Italy. The government's vaccination plan was to administer the vaccine first to healthcare personnel, then to the elderly, then to school and university staff, and ultimately to the rest of the population. However, as the Italian health system is regionally-based, the vaccination plan was affected by a lack of

coordination between the central government and the regional administrations. In addition, the shortage of vaccine doses generated stress in respecting the clauses of the contracts with the EU regarding the planned deliveries. The vaccination plan was lagging behind the initial scheduled timeline and, in March 2021, only 9% of the Italian population received the first vaccine dose.

Figure 1 shows the main pandemic events in both Italy and Israel over time.

Figure 1. Main pandemic events in Italy and Israel

Analysis strategy

We model the multivariate structure in the daily health indicators and the population activity data, referring to major policy decisions (local/national lockdown, reopening). The analysis involves a comparison between two countries which experienced the pandemic as a national threat to their healthcare system: Italy in the first stages (March-April 2020) and Israel during the "second wave" (July-September 2020). The period we are covering is pre-vaccination.

The statistical analysis includes modeling based on various methods designed to enhance information quality, including Bayesian Networks (BN) and Structural Equation Models (SEM). Each modeling approach provides complementary features. The idea is to conduct an ensemble type analysis combining outcomes from various models (Kenett and Salini, 2011). The analysis flows through three sequential steps starting from an initial descriptive BN analysis, aiming at finding the most important relationships with the highest arc strengths between variables. This, in turn, is the basis for a SEM where relationships with the most important latent variables are found. The third stage is a 'what-if scenario' using BN analysis where assumptions of multiple configurations are considered. This is reported in Figure 2. In the final phase, when various scenarios are investigated, the data is discretized using local thresholds in order to enhance interpretability of the scenarios.

Figure 2. Flow chart describing the ensemble analysis

Results

The analysis considers three groups of variables collected daily in Israel and Italy from February 24th, 2020 to January 21st, 2021. These are: (i) mobility variables from Google Community mobility reports; (ii) restriction variables from the University of Oxford's COVID19 government response tracker and (iii) health variables from Israeli and Italian Ministries of Health.

Mobility variables are reported as changes in mobility with respect to a pre-COVID19 baseline. We include mobility changes in retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, and residential. Restriction variables are ordinal variables (from 0 = 'no measures at all' to a maximum of 3 = 'required closing of all events') expressing the intensity of governmental measures to tackle the spread of the virus. These were related to school closing, workplace closing, gathering restrictions, transport closing, stay-at-home restrictions, international movement restrictions, internal movement restrictions. We also consider a stringency index variable ranging from 0 to 100, as a weighted average of the restriction variables.

Health variables include the number of daily COVID19-related deaths, the number of daily hospitalizations due to COVID19 symptoms and the number of daily COVID19 ICU admissions (Savaris et al., 2021). We also consider a variable labeled 'wave' to detect different behavior observed in subperiods of different intensity levels in each country within the period of observation.

For Italy we consider the following subperiods: *Italy's subperiod 1* from February 24th, 2020 to June 15th, 2020, corresponding to the so-called “Italian first pandemic wave” lasting from the very first case detected in Italy to the end of the first lockdown, *Italy's subperiod 2* from June 16th, 2020 to September 15th, 2020, corresponding to the first period of pandemic ease, lasting from the end of the first lockdown to the initial new curbs introduced for the so-called “Italian second pandemic wave”, *Italy's subperiod 3* from September 16th, 2020 to October 8th, 2020, corresponding to the beginning of the “Italian second pandemic wave”, lasting from the introduction of new measures until the introduction of the so-called “color system”, a four-level system for regional restrictions from “white” = “no restrictions” to “red” = “full lockdown”, *Italy's subperiod 4*, from October 9th, 2020 to November 11th, 2020, corresponding to the most acute period of pandemic in Italy, *Italy's subperiod 5*, from November 12th, 2020 to January 21st, 2021, from the ease of the second pandemic wave to the initial vaccination campaign in Italy.

For Israel we consider the following subperiods: *Israel's subperiod 1*, from February 24th, 2020 to March 17th, 2020, since the first case had been detected until the first lockdown, *Israel's subperiod 2*, from March 18th, 2020 to April 26th, 2020, corresponding to the so-called first wave of the pandemic in Israel, where restrictions were imposed to the population, *Israel's subperiod 3*, from June 9th, 2020 to November 3rd, 2020, corresponding to the second pandemic wave in Israel, *Israel's subperiod 4*, from November 4th, 2020 to January 21st, 2021, corresponding to the third pandemic wave in Israel and during which a massive vaccine campaign started.

Bayesian network analysis

The main group of variables of interest is the group of health variables. First, we blacklist in the BN the arcs between variables within each group. Then we create structures using hill-climbing BNs for Israel and Italy by focusing on variables (nodes) with an association to the health variables. Since the effect of mobility and restrictions are observed with delay, after investigating several possible lags, we decided to use 10-day lags for the hospitalizations, 15 days for the number ICUs and 20 days for the number of deaths with respect to the time where measures were implemented. This followed an arc strength analysis of arcs pointing towards the health nodes, looking for the highest average strength among different lags. Figure 3 presents arc strength levels between mobility and restriction nodes vs. the hospitalization node is presented according to different lags. The same plots have been obtained for ICU numbers and deaths, where the more reasonable lags were 15 days and 20 days, respectively.

Figure 3. Strength analysis. Arc strengths according to different lags. Restriction nodes vs. hospitalizations node for Italy (a) and Israel (b). Mobility nodes vs. hospitalizations node for Italy (c) and Israel (d).

Figure 4 reports the derived BNs for Italy and Israel. Measuring the degree of confidence in a graphical structure of a BN is a key problem in inference. Figure 5 reports the same networks, but with the arc strength obtained by applying nonparametric bootstrap to the data and estimating the relative frequency of the feature of interest (Friedman, Goldszmidt and Wyner, 1999). Moreover, Figure 5 (c) and (d) display the node ‘*deaths*’ highlighted together with its parent nodes and the relative arcs with the estimated strength for the networks for Italy and Israel, respectively. All graphs in Figures 4 and 5 can be downloaded in html format from the repository at the link <https://tinyurl.com/rbexhtww> where it is also possible to browse and explore them interactively.

Figure 4 Bayesian Network structure. (a): network for Italy; (b): network for Israel. Both figures are downable and browsable on <https://tinyurl.com/rbexhtww>.

Figure 6. Comparison of lagged BNs for Italy and Israel, Italy is the target BN, Israel is the current BN. Green represents true positive, blue represents false positive and dashed red represents false negative. The resulted Hamming index is.

Structural equation modeling

After identifying the arcs between nodes with the BN, we use SEM to understand which arcs were significant and also to check the robustness of the bootstrap approach to BNs. In the discrete networks that we used for the scenarios, we set whitelists based on the significant arcs in the SEM model. Whitelists are arcs manually imposed on the BN.

In applying SEM, we introduce two latent variables: *behave*, representing the behavioral data and *health*, driving the health-related data (number of death cases with lag of 20 days, number of hospitalized in ICU with lag 15 days, and number of hospitalized patients (with lag of 10 days). Figures 7 and 8 present the SEM path diagram for Italy and Israel. Tables 1 and 3 present the loadings, regressions and covariances for the Italian data and Israeli data. Tables 2 and 4 present the summary of model fit indices for Italy and Israel. Results show that the population behavior was found to be significantly predicted by wave, however, over time, the effect of the "wave" is reduced, as attendance at public places, such as workplaces, transit stations, groceries and pharmacies, and retail and recreation, is not reduced. Also, hospitalization decreased death cases. Restrictions, such as internal lockdowns, international border closings and gathering restrictions led to an increase in health measures. The latent variables, *health* and *behave* were significantly and negatively correlated – when health measures increase, the public reacts by decreasing behavioral activities and vice versa.

Figure 7. Path diagram for the Italian data (JMP Pro version 16)

Regressions	Estimate	SE	Prob> Z
workplaces → [hosp lag10]	0.0051733	0.0013351	0.0001*
workplaces → [death lag20]	-0.001852	0.0012394	0.1350
workplace closing → [icu lag15]	-0.030367	0.0278552	0.2756
workplace closing → [hosp lag10]	0.0667816	0.0270191	0.0134*
wave → workplaces	3.19686	0.5874759	<.0001*
wave → workplace closing	0.0754524	0.0217294	0.0005*
wave → transport closing	0.0837236	0.0126287	<.0001*
wave → transit stations	2.9688032	0.4265175	<.0001*
wave → [icu lag15]	0.1807456	0.0274846	<.0001*
wave → [hosp lag10]	0.2486583	0.0245819	<.0001*
wave → [death lag20]	0.4204896	0.1221507	0.0006*
wave → school closing	-0.316598	0.0169194	<.0001*
wave → retail and recreation	3.7666174	0.59323	<.0001*
wave → residential	-1.129711	0.2063731	<.0001*
wave → international_movement_restrictions	0.0842087	0.0199013	<.0001*
wave → internal_movement_restrictions	0.0676107	0.0158063	<.0001*
wave → grocery and pharmacy	5.10812	0.5543045	<.0001*
wave → gatherings_restrictions	0.3068871	0.0409729	<.0001*
transport closing → [icu lag15]	0.1455176	0.0273859	<.0001*
stringency index → [icu lag15]	0.0067435	0.0038163	0.0772
stay home restrictions → [hosp lag10]	0.0472529	0.022111	0.0326*
[icu lag15] → [death lag20]	1.1659765	0.2939227	<.0001*
[hosp lag10] → [death lag20]	-2.089182	0.410195	<.0001*
retail and recreation → [hosp lag10]	-0.005147	0.0014207	0.0003*
residential → [hosp lag10]	0.013011	0.004218	0.0020*
parks → [hosp lag10]	0.0019993	0.0004296	<.0001*
international_movement_restrictions → [icu lag15]	0.0630446	0.018092	0.0005*
internal_movement_restrictions → [icu lag15]	0.1338373	0.0667171	0.0449*

internal movement restrictions → [hosp lag10]	0.230179	0.0574105	<.0001*
internal movement restrictions → [death lag20]	0.5626452	0.1869919	0.0026*
grocery and pharmacy → [hosp lag10]	0.000544	0.0008651	0.5294
gatherings restrictions → [icu lag15]	-0.063505	0.0407235	0.1189
gatherings restrictions → [hosp lag10]	-0.061278	0.035522	0.0845
gatherings restrictions → [death lag20]	-0.141549	0.0740964	0.0561
Covariances	Estimate	SE	Prob> Z
behave ↔ health	-24.47497	8.5668423	0.0043*
grocery and pharmacy ↔ retail and recreation	87.867134	9.2979598	<.0001*
residential ↔ workplaces	-10.21344	1.4937324	<.0001*
stay_home_restrictions ↔ internal_movement_restrictions	0.7703725	0.0643261	<.0001*
stringency_index ↔ internal_movement_restrictions	8.9560929	0.7747634	<.0001*
stringency_index ↔ stay_home_restrictions	8.8102202	0.7537571	<.0001*
stringency_index ↔ transport_closing	3.708656	0.3479073	<.0001*
workplaces ↔ parks	-275.3571	24.166104	<.0001*

Table 1. Loadings, regressions and covariances for the Italian data

Sample Size	333
Iterations	77
-2 Log Likelihood	22928.544
Number of Parameters	87
AICc	23165.042
BIC	23433.852
ChiSquare	3148.5972
DF	102
Prob>ChiSq	0
CFI	0.7122846
RMSEA	0.2994921
Lower 90%	0.2905408
Upper 90%	0.3085381

Table 2. Summary of fit for the Italian data

Figure 8 presents the SEM path diagram for Israel

Figure 8. Path diagram for the Israeli data (JMP Pro version 16)

Table 3 presents the loadings, regressions and the covariances for the Israeli data. The data from Israel shows direct effect of restrictions on the health outcomes. International movement restrictions decreased the number of hospitalized and death cases. However, it was also found that transport closing decreased ICU cases, but increased hospitalizations. In terms of population behavior, visits in transit stations decreased hospitalizations, while visits in groceries and pharmacies increased it.

Regressions	Estimate	SE	Prob> Z
gatherings restrictions → [death_lag20]	0.0834036	0.0646139	0.1968
grocery and pharmacy → [hosp_lag10]	0.0073164	0.0025899	0.0047*
internal movement restrictions → [death_lag20]	0.2133128	0.0448712	<.0001*
internal_movement_restrictions → [hosp_lag10]	-0.070286	0.0681075	0.3021
internal movement restrictions → [icu_lag15]	0.2298243	0.0530456	<.0001*
international movement restrictions → [death_lag20]	-0.494448	0.06866	<.0001*
international movement restrictions → [hosp_lag10]	-0.550245	0.0422315	<.0001*
residential → [hosp_lag10]	-0.029104	0.0083352	0.0005*
retail and recreation → [hosp_lag10]	0.006643	0.0042186	0.1153
[hosp_lag10] → [death_lag20]	0.3147609	0.0827129	0.0001*
[hosp_lag10] → Lag[icu_lag15]	0.4229745	0.0725468	<.0001*
transit stations → [hosp_lag10]	-0.041554	0.0071265	<.0001*
transit stations → [icu_lag15]	-0.003296	0.002778	0.2354
transport closing → [hosp_lag10]	0.6199998	0.0924994	<.0001*
transport closing → [icu_lag15]	-0.352726	0.0735336	<.0001*
wave → gatherings_restrictions	0.2335669	0.026049	<.0001*
wave → international_movement_restrictions	-0.142389	0.0416318	0.0006*
wave → parks	1.7042613	1.3502963	0.2069
wave → [hosp_lag10]	0.2830623	0.0341744	<.0001*

wave → transport closing	0.0786077	0.0205476	0.0001*
wave → workplace closing	0.1612515	0.0261237	<.0001*
workplace closing → [hosp lag10]	0.3268637	0.1029741	0.0015*
workplace closing → [icu lag15]	0.0525362	0.0701275	0.4538
Covariances	Estimate	SE	Prob> Z
behave ↔ health	1.8928684	0.4624537	<.0001*
gatherings restrictions ↔ behave	-15.63618	1.381421	<.0001*
internal movement restrictions ↔ workplace closing	0.5582769	0.0497297	<.0001*
residential ↔ grocery_and_pharmacy	21.020041	2.8550311	<.0001*
icu lag15] ↔ international movement restrictions	-0.285663	0.0450603	<.0001*
transit stations ↔ transport closing	0.7953274	0.1034623	<.0001*
wave ↔ transit stations	-2.446618	0.3286085	<.0001*

Table 3. Loadings, regressions and covariances for the Israeli data

Table 4 presents measures of fit for the Israeli data.

Sample Size	333
Iterations	11
-2 Log Likelihood	17853.737
Number of Parameters	66
AICc	18018.986
BIC	18237.075
ChiSquare	1035.5668
DF	53
Prob>ChiSq	2.93e-182
CFI	0.8006927
RMSEA	0.2359506
Lower 90%	0.2235195
Upper 90%	0.248608

Table 4. Summary of fit for the Israeli data

The SEM results confirm the significant arcs of the BN. The SEM comparative fit indices (CFI) are relatively low but need to be considered in the context of the overall analysis workflow and not as a stand-alone measure.

'What-if' scenario BN analysis and strength analysis

We now present a “what-if” scenario-based BN analysis where we discretize the health variables and see what happens by setting the categories of some mobility or restriction variables to specific values which of interest. Moreover, inverse inference can be done fixing the levels of the health variables and seeing what happens to the parent nodes. Discrete network whitelists were set up based on the SEM results reported in Tables 1 and 3. This conditioning provides decision makers with a powerful decision support tool used in policy making discussions.

Health variables are discretized according to the cut-offs reported in Figure 9. Cut-offs threshold values were chosen to represent the different health care systems and sizes of the two countries. Figure 10 shows the distribution of daily deaths in Italy (a) and in Israel (b), in terms of number of days distributions in the three classes formed by the cut-off points. The restriction variables are ordinal, so they are not discretized. For the behavior/ activity variables, the Hartemink's algorithm (Hartermink, 2001) is used with a number of cut offs equal to 3.

Figure 9. Cut-off points for health variables for Italy and Israel

Figure 10. Discretization results for COVID19 deaths variables in Italy (a) and in Israel (b). Class limits are reported on the x axes.

Figures 11 and 12 show two scenarios for Italy. In the first scenario, we set evidence as the lowest level of the *international movement restriction* variable. In the second scenario, we set evidence as the highest level of the *international movement restriction* variable.

The distribution of the variable *deaths* changes in the following conditions. When restrictions are present, the number of days when the number of daily deaths is low (less than 100 cases) is 40%. It is 32% when restrictions are not present. The same happens for Israel, as shown in Figures 13 and 14. In Israel, the impact of the *international movement restriction* variable is higher than in Italy; when restrictions are present the number of days with a low number of daily deaths (less than 5) is 40%, compared to 12% when restrictions are not present.

Figure 11. Scenario for Italy when *international movements restriction* is set to 0. International movement restriction node is highlighted in grey.

Figure 12. Scenario for Italy when *international movements restriction* is set to 3. International movement restriction node is highlighted in grey.

Figure 13. Scenario for Israel when *international movements restriction* is set to 2. International movement restriction node is highlighted in grey

Figure 14. Scenario for Israel when *international movements restriction* is set to 4. International movement restriction node is highlighted in grey

In Figures 15 and 16, conditional inference is applied for Italy. With daily deaths level is kept fixed, we observe what happens in the parent nodes. In Figure 16 the supposed scenario is the one where daily deaths are always lower than 100, whereas in Figure 17 the supposed scenario is the one where daily deaths are always larger than 500. We observe, for example, that *stay home restriction* was low (level 1) is present only in 7.5% of days when daily deaths is larger than 500. When daily deaths is always lower than 100, *stay home restriction* was low (level 1) in 44.1% of the days. We also note that *residential mobility* was high (level 3) in 26.3% of the days when daily deaths were larger than 500 and only 12.5% of days when daily deaths were lower than 100.

Some restriction measure variables behave in unexpected directions. For example, in Italy *mobility behavior* changes with no direct reference to deaths. For example, the *internal movement restriction* variable changed from 23.8% at the maximum rate (=2) when deaths were minimum (=1) to 83.3% when deaths were maximum (=3), meaning less deaths imply less movement. This could be explained by the fact that in Italy restriction measures have been in places for long periods, even when deaths increased or decreased. Therefore, movements remained low even with a low number of deaths or high with a large number of deaths.

Figure 15. Scenario for Italy when daily deaths are less than 100 (level 1). Deaths node is highlighted in grey.

Figure 16. Scenario for Italy when daily deaths are larger than 500 (level 3). Deaths node is highlighted in grey.

In Figures 17 and 18, for Israel, inverse inference is applied, daily deaths level is kept fixed, and the model observes what happens in parent nodes. In Figure 17 the scenario is one where daily deaths are always lower than 5; in Figure 18 the scenario is where daily deaths are always larger than 20. We observe, for example, that when *international movement restriction* was high (level 3), 70.9% of days, daily deaths were lower than 5. When daily deaths were larger than 20, in only 27.5% of days *international movement restriction* was high (level 3).

Figure 17. Scenario for Israel when daily deaths are less than 5 (level 1). Deaths node is highlighted in grey.

Figure 18. Scenario for Israel when daily deaths are larger than 20 (level 3). Deaths node is highlighted in grey.

Methods

Rationale and approach to the analysis

The method applied in this work combines official COVID19 health daily data from the Israeli and Italian Ministries of Health, from February 24th, 2020 until January 21st, 2021. Variables included in the study are: number of COVID19 hospitalized patients, number of severe condition patients (in ICUs), and number of COVID19 attributed per day deaths. We also used Google Mobility data (<https://www.google.com/covid19/mobility/>) for Israel and Italy for the same dates. 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. We include also restriction variables from the University of Oxford's COVID19 government response tracker research group (<https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker>).

To conduct the analysis, we implement a workflow combining several multivariate methods including structural equation models (SEM) and Bayesian networks (BN). We use BN analysis to map the relationships between epidemiological and behavioral variables. Death statistics are used to assess the impact of the pandemic and the number of infected people, as reported on a daily basis. Our analysis proceeds through the following steps.

- First, we perform an initial explorative BN for both Israel and Italy in order to describe the relationships among variables. In this step we also perform a between-node strength analysis to derive a hierarchy of the most important variables affecting ICU, deaths and hospitalization.
- Then we identify from the previous BN the most significant nodes and use these nodes in a SEM to study the significance of the relations among nodes.
- Finally, we built up a 'what-if scenario' BN analysis to model the most epidemiological variables, i.e. ICU, death and hospitalization, in order to analyze the dynamics of the pandemic with respect to behavioral variables in the two countries.

To conduct the analyses we used the package R (www.R-project.org/), and its libraries *bnlearn* and *bnviewer* for the Bayesian Networks, and JMP Pro 16 (www.jmp.com) for the Structural Equation Modeling. We provide below a brief introduction to BN and SEM pointing to references for more details.

Bayesian networks

Formally, BNs are direct acyclic graphs (DAG) whose nodes represent random variables in the Bayesian sense: they can be observable quantities, latent variables, unknown parameters or hypotheses (see Scutari and Denis 2021). The arcs represent conditions of dependence; the nodes that are not connected represent variables that are conditionally independent of each other. Each node is associated with a probability function which takes as input a particular set of values for the variables of the parent node and returns the probability of the variable represented by the node. There are efficient algorithms that perform inference and learning starting from BNs. A BN enables the effective representation and computation of a joint probability distribution (JPD) over a set of random variables. The DAG's structure is defined by a set of nodes, representing random variables and plotted by labeled circles; and a set of arcs representing direct dependencies among the variables and plotted by arrows. Although the arrows represent direct causal connections between the variables, under some conditions, the reasoning process can operate on a BN by propagating information in any direction. A BN reflects a simple conditional independence statement, namely that each variable, given the state of its parents, is independent of its non- descendants in the graph. This property is used to reduce, sometimes significantly, the number of parameters that are required to characterize the JPD. This reduction provides an efficient method to compute the posterior probabilities given the evidence in the data (Kienett, 2019).

Structural equation modeling

Gupta and Kim (2008) propose linking Bayesian networks (BN) to Structural equation modeling (SEM), which has an advantage in testing causal relationships between factors. The capability of SEM in empirical validation, combined with the prediction and diagnosis capabilities of Bayesian modeling, facilitates effective decision making from identification of causal relationships to decision support. SEM is a set of statistical techniques used to measure and analyze the relationships of observed and latent variables (see Hoyle, 1995). It examines linear causal relationships among variables, while simultaneously accounting for measurement error. SEM is widely used in the social sciences and psychology. It provides a flexible framework for developing and analyzing complex relationships among multiple variables that allow researchers to test the validity of a theory using

empirical models. A discussion about SEM and its connection with causal models is reported in Pearl (1998). An interesting application that combines SEM and BN is Yoo and Oh (2013).

Discussion

In this article we present a methodology to examine the mutual effect of national policy and population behavior. This can serve as an effective tool for policy makers evaluating anticipated pandemic mitigation policies. Such assessments are dependent on health conditions, desired outcomes and variables such as population's compliance over time and national and local health systems' capacity. We apply a multi-methods approach combining Bayesian networks analysis, structural equation modeling and 'what-if' Bayesian networks scenarios to examine the interrelations of mitigation policies, population behavior and health indicators.

Given that COVID19 is a human to human transmitted disease, with high infection rates, pandemic mitigation policies should limit people physical interactions, but at the same time, allow for maximal continuity of economic activity. As a consequence, managing and applying movement restrictions should take into consideration three factors: (a) the type of the restriction (local, national, the activity being restricted); (b) the duration of the restriction; and (c) the severity of the morbidity. Gaining compliance and adherence to the restriction policies depends on these three factors. As the results in our analysis show, over time, the population compliance decreased in both countries, even when health indicators were not satisfying. This highlights the importance of applying a selective and differential approach in various public health conditions. For example, restricting international movement had a larger effect on the number of death cases. At the same time, other restrictions, such as in workplaces and transport stations, had a moderate effect on the number of hospitalized patients. Also, behavioral changes, such as attendance at grocery stores and pharmacies, workplace and transit stations, had a very small effect on health indicators. This means that people limit attendance at public places according to their personal risk assessment, while the consequences of such limitations vary according to the duration of the restrictions and their type. For example, we found that in Israel, higher levels of deaths cases are followed by internal restrictions, i.e. lockdowns, but also higher attendance at public place, such as retail stores and transit stations. In Italy, higher death rates are the result of less restrictions, and higher levels of hospitalized patients.

The analysis also highlights the strong effect of internal and international travel restrictions on COVID19 deaths rates. This is in comparison to other restrictions which have a smaller effect on the number of hospitalized patients. This finding is important as vaccinations and improved healthcare systems provide effective solutions that permit lower levels of restrictions and lower levels of population compliance with instructions and limitations. More importantly, it allows a country to maintain its economic activity running with acceptable levels of morbidity and hospitalizations. However, it seems that international restrictions, such as airport closures, have a strong and direct effect on the number of deaths. In terms of mitigation policy, closing international borders appear as an effective tool for a quick reduction of death cases, when the pandemic exceeds acceptable limits.

Although critical in the pandemic spread, the population behavior was found to have only a moderate direct effect on the health indicators. In fact, both countries have reacted in an opposite way at early stages of the pandemic. On February 2020, Italy was late in reacting, with no restrictions and extreme hospitalizations, ICU and death cases. In Israel, immediately after detecting the first cases, a national total lockdown was applied, which led to low levels of morbidity for few months. However, later, both countries experienced higher levels of morbidity with restrictions, but with lower levels of public's compliance. This is important for policy decision making: considerations. Future pandemic management should assess the public's ability and willingness to comply with the directions. Apparently, the severity of the health indicators does not guarantee full compliance with the instructions. The effect of restrictions while the public does not comply is dramatic as reflected by higher levels of morbidity and death rates.

Dealing with a pandemic like COVID19 needs to account for possible multiple strategies. The continuous relationship over time among the variables involved, renders an approach with BNs and SEM particularly suitable. The reversed conditional analysis described above is useful to "learn from experience". Restriction measures can be successful or not. This is why we model a variable "wave" to temporally and spatially contextualize the restriction measures following, or preceding, the introduction of restriction measures and consequently affecting the behavior of the population. The current analysis has several limitations. In general, we show the aggregated behavior of variables with respect to key healthcare variables (hospitalization, ICU and deaths).

Direct consequences of a measure put in place, at a certain time t , should be studied with a time series model. We construct scenarios regarding choices taken to overcome pandemic problems. These are summarized by counterfactual questions such as: "if in this situation one would have done this, what would have happened to this quantity?". The COVID19 pandemic is a continuous event that needs to be managed as such. The definition of a "wave", in each country, reflects increased levels of health indicators during a certain period. Health data observatories and surveillance systems, across countries, need to calibrate the data and provide a multivariate perspective. In this paper we provide examples of both.

References

- Ashkenazi, I. and Rapaport, C. (2020) Saving Lives Versus Saving Dollars: The Acceptable Loss for Coronavirus Disease 2019. *Critical Care Medicine*, 48(8): 1243-1244.
- 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
- Bargain, O., & Aminjonov, U. (2020). Trust and compliance to public health policies in times of COVID-19. *Journal of Public Economics*, 192, 104316.
- 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
- Bodas, M., & Peleg, K. (2020). Income assurances are a crucial factor in determining public compliance with self-isolation regulations during the COVID-19 outbreak—cohort study in Israel. *Israel Journal of Health Policy Research*, 9(1), 1-10.
- Borgonovi, F., & Andrieu, E. (2020). Bowling together by bowling alone: Social capital and Covid-19. *Social Science & Medicine*, 265, 113501.
- Bonny-Noach, H., & Gold, D. (2020). Addictive behaviors and craving during the COVID-19 pandemic of people who have recovered from substance use disorder. *Journal of addictive diseases*, 1-20.
- Bradbury-Jones, C. & Isham, L. (2020). The pandemic paradox: the consequences of COVID-19 on domestic violence. *Journal of Clinical Nursing*, 29(13-14):2047-2049
- Chernozhukov, V., Kasahara, H., & Schrimpf, P. (2021). Causal impact of masks, policies, behavior on early covid-19 pandemic in the US. *Journal of econometrics*, 220(1), 23-62.
- Calina, D., Hartung, T., Mardare, I., Mitroi, M., Poulas, K., Tsatsakis, A., ... & Docea, A. O. (2021). COVID-19 pandemic and alcohol consumption: Impacts and interconnections. *Toxicology Reports*.
- Czeisler, M. É., Lane, R. I., Petrosky, E., Wiley, J. F., Christensen, A., Njai, R., ... & Rajaratnam, S. M. (2020). Mental health, substance use, and suicidal ideation during the COVID-19 pandemic—United States, June 24–30, 2020. *Morbidity and Mortality Weekly Report*, 69(32), 1049-1057.
- Dryhurst, S., Schneider, C. R., Kerr, J., Freeman, A. L., Recchia, G., Van Der Bles, A. M., ... & van der Linden, S. (2020). Risk perceptions of COVID-19 around the world. *Journal of Risk Research*, 23(7-8), 994-1006.
- Engle, S., J. Stromme, and A. Zhou (2020): Staying at home: Mobility effects of Covid-19, Discussion paper, Covid Economics: Vetted and Real Time Papers, available at SSRN: <http://dx.doi.org/10.2139/ssrn.3565703>.
- Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., ... & Bhatt, S. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 584(7820), 257-261.
- Friedman, N., Goldszmidt, M., & Wyner, A. J. (1999). On the application of the bootstrap for computing confidence measures on features of induced Bayesian networks. In *AISTATS*.
- Gupta, S., & Kim, H. W. (2008). Linking structural equation modeling to Bayesian networks: Decision support for customer retention in virtual communities. *European Journal of Operational Research*, 190(3), 818-833.
- Hoyle, R. H. (1995). *Structural equation modeling: Concepts, issues, and applications*. Sage.
- Jordan, J., Yoeli, E., & Rand, D. (2020). Don't get it or don't spread it? Comparing self-interested versus prosocially framed COVID-19 prevention messaging. Pre-print, <https://doi.org/10.31234/osf.io/yuq7x>
- Kawohl, W., & Nordt, C. (2020). COVID-19, unemployment, and suicide. *The Lancet Psychiatry*, 7(5), 389-390

- Kenett, R.S., Salini, S. (2011). Modern analysis of customer satisfaction surveys: comparison of models and integrated analysis. *Applied Stochastic Models in Business and Industry*, 27(5): 465--475.
- Kenett, R.S. (2019) Applications of Bayesian Networks *Trans. on Machine Learning and Data Mining*, Vol.12, No 2, pp. 33-52. http://www.ibai-publishing.org/journal/issue_mldm/2019_October/12_2_33_54_mldm.pdf
- 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.
- Liu, S., Heinzel, S., Haucke, M. N., & Heinz, A. (2021). Increased Psychological Distress, Loneliness, and Unemployment in the Spread of COVID-19 over 6 Months in Germany. *Medicina*, 57(1), 53.
- 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)
- Painter, M., & Qiu, T. (2021). Political beliefs affect compliance with government mandates. *Journal of Economic Behavior & Organization*, 185, 688-701.
- Pearl, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods & Research*, 27(2), 226-284.
- Pfefferbaum, B., & North, C. S. (2020). Mental health and the Covid-19 pandemic. *New England Journal of Medicine*, 383(6), 510-512.
- Plohl, N., & Musil, B. (2021). Modeling compliance with COVID-19 prevention guidelines: The critical role of trust in science. *Psychology, Health & Medicine*, 26(1), 1-12.
- Savaris, R.F., Pumi, G., Dalzochio, J. et al. (2021). Stay-at-home policy is a case of exception fallacy: an internet-based ecological study. *Scientific Reports*, 11: 5313.
- Scutari, M., & Denis, J. B. (2021). *Bayesian networks: with examples in R*. CRC press.
- Wang, D., Marmo-Roman, S., Krase, K., & Phanord, L. (2021). Compliance with preventative measures during the COVID-19 pandemic in the USA and Canada: Results from an online survey. *Social Work in Health Care*, 60(3), 240-255.
- Weihsuch A. Chiu, Rebecca Fischer, Martial L. Ndeffo-Mbah, State-level needs for social distancing and contact tracing to contain COVID-19 in the United States, *Nature Human Behaviour*, 10.1038/s41562-020-00969-7, 4, 10, (1080-1090), (2020).
- Xue, B., & McMunn, A. (2021). Gender differences in unpaid care work and psychological distress in the UK Covid-19 lockdown. *PloS one*, 16(3), e0247959
- Yan, Y., Malik, A. A., Bayham, J., Fenichel, E. P., Couzens, C., & Omer, S. B. (2021). Measuring voluntary and policy-induced social distancing behavior during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, 118(16), e2008814118.
- 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, 210-220.
- Yoo, C., & Oh, S. (2013). Combining Structure Equation Model with Bayesian Networks for predicting with high accuracy of recommending surgery for better survival in Benign prostatic hyperplasia patients. In *20th International Congress on Modelling and Simulation-Adapting to Change: The Multiple Roles of Modelling, MODSIM 2013-Held jointly with the 22nd National Conference of the Australian Society for Operations Research, ASOR 2013 and the DSTO led Defence Operations Research Symposium, DORS 2013* (pp. 2029-2033). Modelling and Simulation Society of Australia and New Zealand Inc.(MSSANZ).