

# The First Wave of COVID-19 in Israel – Initial Analysis of Publicly Available Data

Mark Last<sup>1\*</sup>,

**1** Department of Software and Information Systems Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

\* [mlast@bgu.ac.il](mailto:mlast@bgu.ac.il)

## Abstract

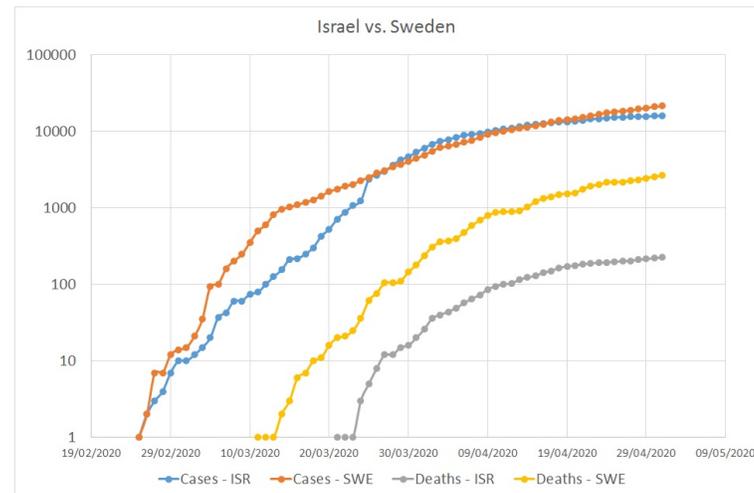
The first case of COVID-19 was confirmed in Israel on February 21, 2020. Within approximately 30 days, the total number of confirmed cases climbed up to 1,000, accompanied by a doubling period of less than 3 days. About one week later, after this number exceeded 4,000 cases, and following some extremely strict measures taken by the Israeli government, the daily detection rate started a sharp decrease from the peak value of 1,131 down to slightly more than 100 new confirmed cases on April 30. Motivated by this encouraging data, similar to the trends observed in many other countries, along with the growing economic pressures, the Israeli government started relaxing its emergency regulations as part of an "exit strategy". This article attempts to analyze the currently available data on Israel and a country of similar size (Sweden) in order to understand the local dynamics of COVID-19, assess the effect of the implemented intervention measures, and discuss some plausible scenarios for the foreseeable future.

## Introduction

The first case of COVID-19 was confirmed in Israel on February 21, 2020. On March 9, a two-week self-isolation was imposed on all people coming from abroad and on March 12, schools and universities were closed by the government order, partially switching to distant teaching. Then, on March 19, when the number of daily confirmed cases exceeded 100, all non-essential businesses were ordered to close, the employees were required to increase social distancing of their workers and, if possible, allow them to work from home, whereas people's movement outside their homes was restricted significantly. Still, by March 23, the number of confirmed cases climbed up to 1,000, accompanied by a doubling period of less than 3 days. A continuous decrease in the growth rate has started only around March 29, after the total number of confirmed cases exceeded 4,000. By April 30, the doubling period went up to several weeks with only 222 death outcomes out of 15,946 confirmed cases.

Sweden, a country slightly more populated than Israel, detected its first COVID-19 cases around February 26. The evolution of the epidemic in both countries is compared in Fig 1. Initially, the total number of confirmed cases grew up much faster in Sweden than in Israel, exceeding 1,000 on March 15. However, despite much softer intervention policy in Sweden than in Israel, both countries kept reporting similar amounts of detected cases from the end of March until the second half of April, when Sweden started to surpass Israel, reaching a gap of 5,000 cases by the end of April. However for several critical weeks, the significant policy differences between the two countries

seemed to have a limited effect on their reported infection rate. On the other hand, Sweden is suffering a much higher mortality rate than Israel, with the current ratio of about 10 : 1 between the amount of death outcomes in the two countries.



**Fig 1. Comparing Israel to Sweden.** Confirmed cases and death outcomes

With the peak unemployment rate of 24%, mostly due to people placed on unpaid leave as a result of the pandemic, the Israeli government decided to partially lift the lockdown, which has been maintained throughout the Passover holiday (April 7-16). The "exit strategy" steps under consideration included increasing the allowed percentage of workers at workplaces, opening some shops, partially resuming public transportation, gradually getting children back to school, and selectively releasing people from the movement restrictions based on the level of infection in their neighborhood and their risk group. A list of detailed measures for easing the current restrictions was approved by the Israeli government on May 04, 2020. The ongoing discussions of alternative "exit strategies" brings up the following questions:

1. What is the basic reproduction number  $R_0$  in Israel vs. Sweden and how is it evolving over time?
2. Which measures, if any, have caused the steep decrease in the growth rate of confirmed cases in Israel, which has been observed since the end of March 2020 ?
3. What is the true Infection Rate (IR) in Israel, i.e. what is the actual fraction of infected people including those who were not tested for coronavirus during their detectable period ?
4. What are the Case Fatality Rate (CFR) and the Infection Fatality Rate (IFR), defined as the ratio of COVID-19 attributed deaths to the number of confirmed cases and the total number of infection cases, respectively.
5. Finally, the most significant question: What will be the further evolution of the epidemic in Israel in the view of the prospective exit strategy?

The subsequent sections will focus on each one of these crucial questions. Finally, we will discuss the implications of our data analysis on the prospective decisions to be taken by the government in Israel and other corona-affected countries.

## Materials and methods

### Data Sources

For daily case count data in Israel and other countries, we rely upon the Novel Coronavirus (COVID-19) Cases Data, which is part of Humanitarian Data Exchange (HDX) platform [1]. The data is compiled since 22 January 2020 by the Johns Hopkins University Center for Systems Science and Engineering (JHU CCSE) from various sources such as the World Health Organization (WHO) and European Centre for Disease Prevention and Control (ECDC). The available fields include Province/State, Country/Region, Last Update, Confirmed, Suspected, Recovered, and Deaths.

The coronavirus testing data in Israel was downloaded from the COVID-19 Data Repository of the Israeli Ministry of Health [2] on April 28, 2020. The available fields included: test\_date, binary symptom indicators (cough, fever, sore\_throat, shortness\_of\_breath, head\_ache), corona\_result (positive, negative or other), binary age indicator (age\_60\_and\_above), gender, and test\_indication (Abroad, Contact with confirmed or Other). The data file contained the results of 313, 135 tests performed between March 11 and April 25, 2020.

### Modeling the Infection Dynamics

Following [5], we assume that the COVID-19 outbreak can be represented by SIR dynamics, which assumes that at any given point of time, each individual in the population belongs to one of the following three states: (I)nfectious, (S)usceptible to infection, or (R)emoved from the transmission process. However, the original SIR model and its variations, like SEIR (susceptible–exposed–infected–recovered) [4], are continuous time models based on a set of differential equations. In contrast, we represent the COVID-19 dynamics by a discrete time model, which is more appropriate for the time series data of daily case counts. Fixing the total size of the population under a discrete time SIR model implies that the sum of the daily changes in the amount of individuals belonging to each one of the above three compartments should be equal to zero:

$$I_t + S_t + R_t = 0 \quad (1)$$

where  $I_t$ ,  $S_t$ , and  $R_t$  stand for the changes in the total amount of infected, susceptible, and recovered people, respectively, between the days  $t - 1$  and  $t$ . Our discrete time model for the temporal evolution of the number of infected individuals builds upon the basic reproduction number  $R_0$ , which represents the average number of secondary infections an infected person will cause before he or she is effectively removed from the population as a result of recovery, hospitalization, quarantine, etc. [5]. As long as  $R_0 > 1$ , the number of infected people will grow exponentially. At the time of a pandemic, the value of  $R_0$  may be reduced by decreasing the amount of social interaction (a measure known as "social distancing"). Additional parameters used by our model include the average duration of the incubation period  $L_I$  between the exposure and the onset of clinical symptoms and the generation time  $[L_I - max.T_I; L_I - min.T_I]$  between the primary case and the secondary case, where  $min.T_I$  and  $max.T_I$  represent, respectively, the minimum and the maximum number of days from the start of the infectious period to the onset of the symptoms. We assume that symptomatic individuals are not infectious anymore as they are immediately isolated from the susceptible population and required to take a COVID-19 test. We also assume that asymptomatic people are never referred to a test and that the social behavior of infected individuals does not change during their infectious period. Hence,

we can define the following discrete time model of infection dynamics:

$$I_t = \frac{R_0}{\max T_I - \min T_I + 1} \sum_{i=t-L_I+\min T_I-1}^{t-L_I+\max T_I-1} I_i \left(1 - \frac{\sum_{j=1}^i I_j}{N}\right) \quad (2)$$

where  $N$  is the initial size of the susceptible population. Given a time series of estimated daily values  $I_i$ , the incubation period  $L_I$ , the generation time range  $[L_I - \max T_I; L_I - \min T_I]$ , and the population size  $N$ , we can use a stochastic optimization method to find the value of  $R_0$ , which should reconstruct that series with a minimal average absolute error. The error can be further reduced by splitting the time series into several segments and estimating  $R_0$  for each segment separately.

To monitor the evolution of  $R_0$  on daily basis, we can take several simplifying assumptions. First, we assume that as long as only a small proportion of a country's population is infected, the number of susceptible individuals  $S_t$  remains close to the total population size  $N$ . We also assume the generation time to be fixed to the midpoint day of the interval  $[L_I - \max T_I; L_I - \min T_I]$ . Applying these assumptions to Eq 2, we can approximate the expected number of new infected cases on day  $t$  as:

$$I_t = R_0 I_{t-L_I+(\min T_I+\max T_I)/2-1} \quad (3)$$

Eq (3) is a linear autoregressive process of order  $L_I - (\min T_I + \max T_I)/2 + 1$  and thus we can use the least squares method to estimate the value of  $R_0$  from the daily counts of new infected cases. Since  $R_0$  may be affected by multiple factors, such as the level of social distancing, we can monitor its change over time by calculating the moving slope of the autoregression equation (3) over a sliding window of  $n$  days.

To project the epidemic dynamics into the future, we just need to choose the expected value of  $R_0$ , initialize the daily amounts of infected people on the last  $L_I$  days before the beginning of the projection period, and start calculating the succeeding values of  $I_t$  recursively using Eq (2). A similar simulation paradigm was implemented in [7] to evaluate several transmission and intervention scenarios for COVID-19 in the United States over the next five years.

As long as there is no daily testing of the entire population, the exact values of  $I_t$  remain unknown. However, in the next sub-section, we explain how one can estimate the actual dynamics of infected cases from the currently available epidemiological data.

## Estimating the Actual Infection Dynamics from Available Data

We assume that symptomatic individuals take a COVID-19 test on the day of the symptoms onset (the end point of the incubation period), but it takes  $r$  days on average to report a positive test result as a "confirmed case". Given the average incubation period  $L_I$  and the average reporting delay  $r$ , the time series of the daily number of new infection cases  $I_t$  is related to the time series of the daily number of confirmed cases  $C_t$  by Eq (4).

$$I_t = p_{t+L_I+r} C_{t+L_I+r} \quad (4)$$

where  $p_t$  is the reporting rate on day  $t$ , which is equal to 1 if there is a full and timely detection of all positive cases, which requires daily testing of the entire population, and greater than 1 when some under-reporting takes place.

While the amount of new confirmed COVID-19 cases in various countries, regions and even cities is published on daily basis, its level of under-reporting is unknown and potentially unstable due to inconsistent testing policy, varying reporting time, and other factors and thus we cannot use it as a trustworthy indicator of the epidemic dynamics. On the other hand, the daily number of virus-related deaths  $D_t$  is expected to be much

more reliable than the testing results and it should not suffer from any significant reporting delay. Assuming a fixed Infection Fatality Rate  $IFR$ , we can estimate the daily amount of infection cases on day  $t$  as a function of the amount of death outcomes on day  $t + L_I + d$ , where  $d$  is the average number of days between the onset of the virus symptoms and the patient death (see Eq 5). Hence, disregarding the true value of  $IFR$ , we can explore the dynamics of the daily death rate  $D_t$  as a time-shifted substitute for the true infection rate  $IR_t$  (Eq 6). Given the average reporting delay  $r$ , we can also estimate the Case Fatality Rate (CFR), which is defined in epidemiology as the proportion of people who die from a specified disease among all individuals diagnosed with the disease over a certain period of time [6]. Since the CFR numerator is restricted to deaths among people included in the denominator, we calculate it as the average ratio between the amount of death outcomes on day  $t + d - r$  and the number of confirmed cases on day  $t$  (Eq 7).

$$I_t = \frac{D_{t+L_I+d}}{IFR} \quad (5)$$

$$IR_t = \frac{\sum_{i=1}^t I_i}{N} \quad (6)$$

$$CFR = \frac{D_{t+d-r}}{C_t} \quad (7)$$

## Results

### The Epidemic Dynamics in Israel and Sweden

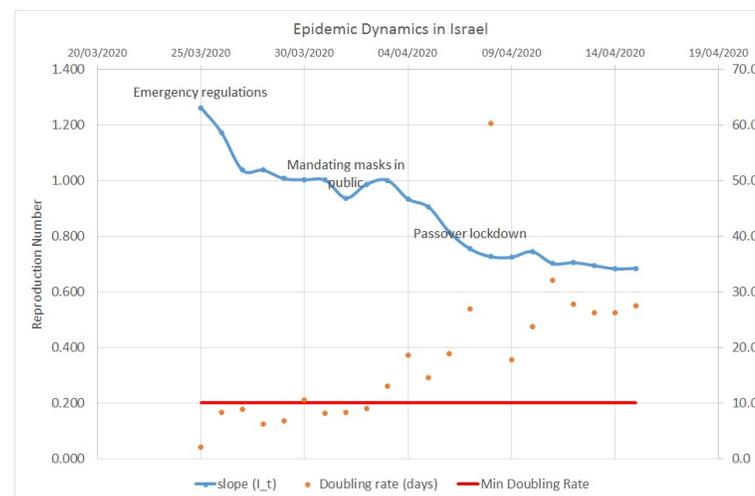
Our analysis of COVID-19 evolution in Israel is based on the following timing parameters provided in [8, 9]:

- Mean incubation period  $L_I = 5.5$  days
- Minimum time from the start of the infectious period to the onset of the symptoms  $min.T_I = 1$  day
- Maximum time from the start of the infectious period to the onset of the symptoms  $max.T_I = 3$  days
- Mean generation period (average number of days it takes a patient to infect others)  $L_I - (min.T_I + max.T_I)/2 = 3.5$  days
- Mean time between the onset of the virus symptoms and the patient death  $d = 11$  days
- Mean time from infection to death  $L_I + d = 16.5$  days
- Mean time from symptoms to detection  $r = 4.5$  days
- Mean time from infection to detection  $L_I + r = 10$  days
- Mean time from detection to death  $d - r = 6.5$  days

Fig 2 shows the evolution of the reproduction number  $R_0$  in Israel between March 25 and April 15, 2020. The daily number of infection cases was estimated by Eq 5, where we used Case Fatality Ratio (CFR) of 1.46% (calculated below) as an estimator of Infection Fatality Ratio (IFR). On each day,  $R_0$  was calculated as a moving slope of a linear autoregressive model over the last 14 days using a lag of 4 days, which is equal to

the mean generation interval. The chart also shows the main intervention measures taken by the Israeli government during the same period. Since we did not have enough data to estimate the infection rate before March 25, we could not evaluate the direct effect of the school closing around March 12 as well as business activity and movement restrictions imposed on March 19.

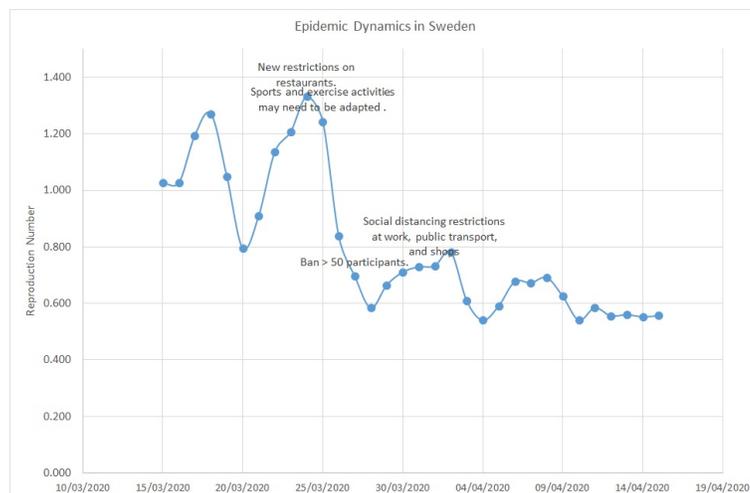
Apparently, the emergency regulations declared on March 25, which forced most Israelis to stay at their homes, have contributed to social distancing by quickly reducing the mean value of  $R_0$  from 1.26 to 1.00 within 5 days only. An additional decrease in  $R_0$  observed in the beginning of April could be caused by more people wearing masks in public. The Passover lockdown is the most likely factor that brought the reproduction number below 0.70 by April 15, which was the last official day of the Holiday. As can be seen on the same chart, after  $R_0$  went below the value of 1.0, the doubling rate (number of days required to double the number of infected people) climbed highly above the 10 days threshold, which was defined in the Government decision from May 04, 2020 as one of the criteria for rolling back the "exit strategy" steps.



**Fig 2. COVID-19 Dynamics in Israel.** March 25 - April 15, 2020

Fig 3 shows the 14-day moving average of  $R_0$  in Sweden from March 15 to April 15, 2020. The daily number of infection cases was calculated using the estimated Infection Fatality Ratio (IFR) of 14.82%. The chart also shows the main intervention measures taken by the Swedish government during the same period. On March 24, the mean reproduction number in Sweden reached its peak value of 1.33, just slightly higher than  $R_0$  observed at that time in Israel (1.26). Following the restrictions on various recreation activities, it went down sharply to 0.58 within just four days. It is noteworthy that in Israel, much stricter restrictions reduced  $R_0$  to 1.0 only. After March 28, the Swedish reproduction number fluctuated between 0.6 and 0.8, eventually converging to 0.56, a lower value than in Israel during the same period (0.68). While it is evident that Sweden did a better job than Israel in terms of social distancing, the effect of various measures on the actual public behavior in each country requires further investigation.

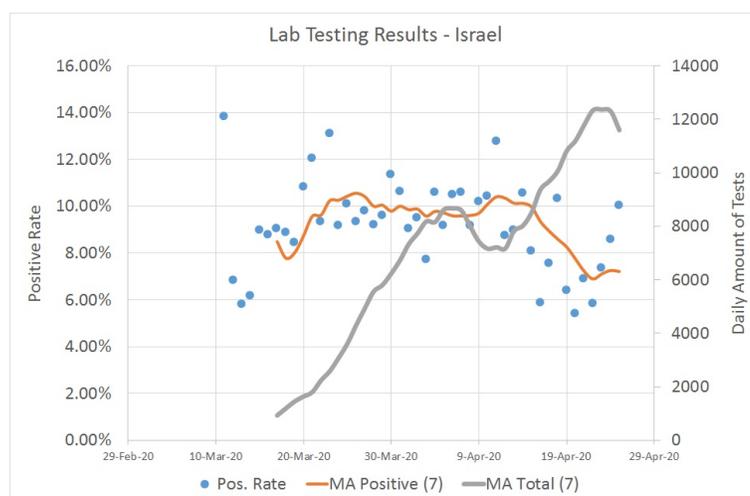
The true Infection Rate (IR) of COVID-19, like any other epidemic, is a non-decreasing function of time. Hopefully, it will be measured in the near future using massive serological tests but its true past values will remain unknown forever. However, we believe, like the authors of [10], that the Test Positivity Rate (TPR), the fraction of positive to total tests, should be an upper bound on the IR during the first few weeks of infection. Fig 4 shows the 7-day moving averages for TPR and for the daily amount of tests between March 11 and April 25. Since the mean TPR reached its peak value of



**Fig 3. COVID-19 Dynamics in Sweden.** March 15 - April 15, 2020

10% around March 25 when the infection rate was also at its maximum, we can conclude that the true IR in Israeli population also did not exceed 10% at that time. This upper bound is almost 100 times higher than the lower bound of 0.10% based on about 10,000 confirmed cases reported by the end of March out of the country's population of nearly 8.9 million.

207  
208  
209  
210  
211

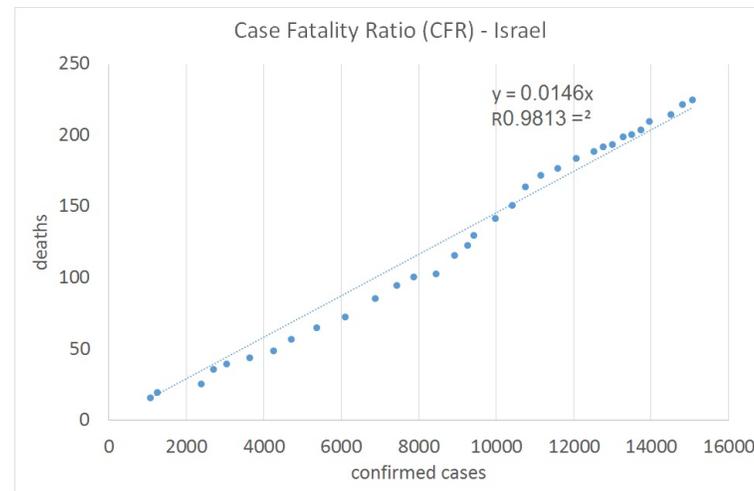


**Fig 4. Lab Testing Results in Israel.** March 11 - April 25, 2020. Source: <https://data.gov.il/dataset/covid-19>

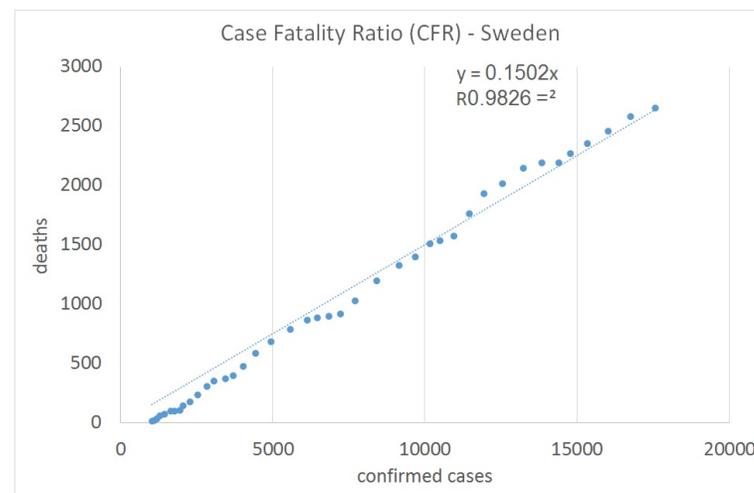
Figs 5 and 6 show the calculation of the Case Fatality Ratio (CFR) in Israel and Sweden, respectively, as a slope of a linear regression model between the cumulative amount of confirmed cases and the cumulative number of death outcomes 7 days later. The line equations, with the intercept set to 0, are  $y = 0.0146x$  for Israel and  $y = 0.1502x$  for Sweden, which implies an order of magnitude difference between the CFR of the two countries (1.46% vs 15%). The high values of R-Square (0.98) in both regression models may indicate that most COVID-19 deaths are related to previously confirmed cases. This result supports the hypothesis that most symptomatic patients are detected by the lab tests and thus the actual Infection Rate should be close to its lower bound. However, we do not have and may never have enough data to test this

212  
213  
214  
215  
216  
217  
218  
219  
220  
221

claim in either country. We also have no plausible explanation for the exceptionally low CFR values in Israel compared to Sweden and to the world average of 3.6% in general. The reasons for the significant gaps between mortality rates in different countries and regions deserve a separate study as they may include differences in lab testing strategies, clinical treatment practices, outcome classification policies, coronavirus strains, etc.



**Fig 5. Case Fatality Ratio in Israel.** Lag: 7 days



**Fig 6. Case Fatality Ratio in Sweden.** Lag: 7 days

## Projecting into the Future

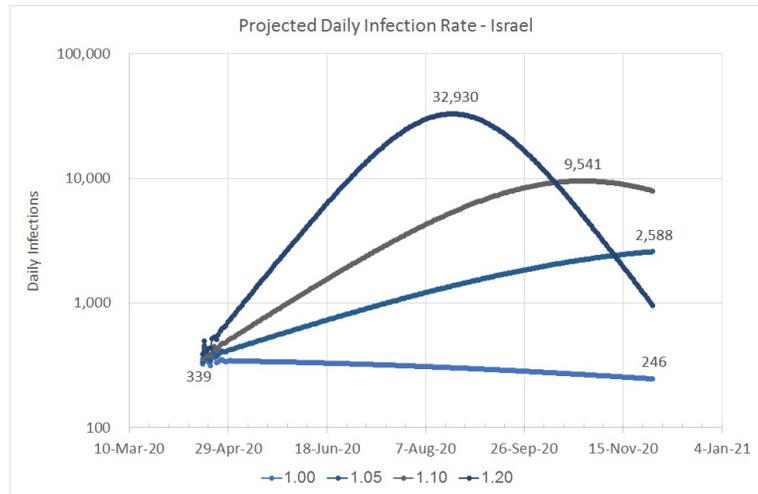
Based on the infection data estimated up to April 15, Figs 7 and 8 simulate the projected number of daily and cumulative infection cases, respectively, for various possible values of the reproduction number  $R_0$ . Our simulation covers the period from the beginning of May to the end of November, when we may be close to a start of a new flu season in Israel. These projected numbers allow us to forecast three important parameters: maximum required number of respirators, total amount of death outcomes, and the extent of herd immunity at the end of the simulated period. Following [9], we assume that the average stay on respirator is equal to 10 days and that the average

proportion of patients needing respirators in Israel is about 2.3%. Consequently, the maximum number of required respirators should be equal to 23% of the maximum number of daily infections. The number of respirators used by COVID-19 patients (70 – 80) in the beginning of May 2020 is very close to the above estimator based on the moving average of new infections started about two weeks earlier. The projected extent of population immunity is calculated as a ratio between the cumulative number of projected infections and the size of the Israeli population (8.9 million), under a conservative assumption that the current Case Fatality Ratio of 1.46% is close to the actual Infection Ratio (IR). The total number of death outcomes can be estimated as a product of the currently known Case Fatality Ratio (1.46%) by the cumulative number of infections.

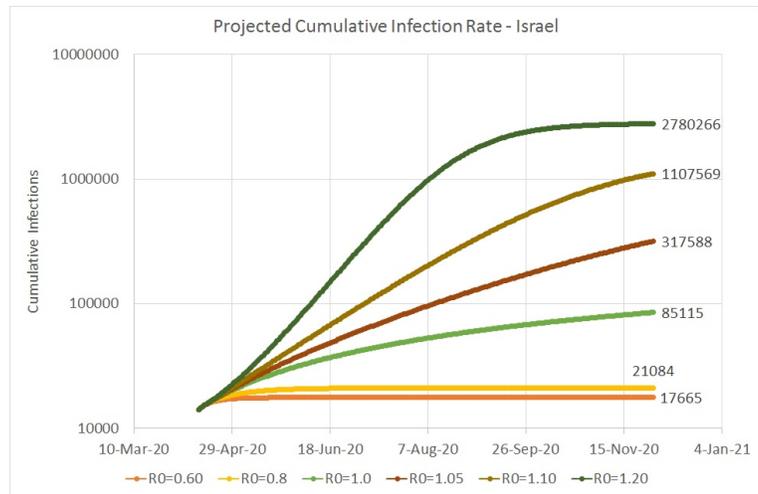
Here is a brief discussion of each simulated scenario:

- $R_0 = 0.60$ . Based on the recent dynamics, where  $R_0$  stayed between 0.6 and 0.7 at the time of a massive lockdown, this value may imply a continuous slowdown of the Israeli economy. However, it is expected to eliminate the local epidemic by the beginning of June at a minimal cost of about 30 additional death outcomes. On the other hand, a vast majority of Israeli population will remain susceptible to the virus, even if the current Infection Rate in Israel is close to its testing-based upper bound of 10%. It is noteworthy that the daily amounts of new cases reported in the beginning of May are very close to the numbers predicted using this value of  $R_0$ .
- $R_0 = 0.80$ . This value was observed in Israel after significant economic restrictions but before the Passover lockdown. If the corresponding amount of social distancing can be maintained by the entire population, the local virus cases should be eliminated by the beginning of August, again without immunity for at least 90% of people. At these and lower values of  $R_0$ , the Government criterion of less than 100 new daily cases from an unknown source should be met most of the time.
- $R_0 = 1.00$ . This value seems to be achievable with minimal economic restrictions. Rather than eliminating the virus, it will keep the daily infection rate at 200 to 300 cases (exceeding the Government criterion of maximum 100 cases), resulting in about 1,000 additional deaths and at least 1% immunity by the end of November. This scenario will require about 70 respirators on average, similar to the number occupied in the beginning of May 2020.
- $R_0 = 1.05$ . This value may be slightly easier to implement in terms of social distancing than  $R_0 = 1.0$ , but it will cause a steady, though a slow growth in the daily infection rate up to 2,500 cases per day by the end of November. This means Israel will need about 600 respirators at that time. According to this scenario, at least 4% of Israeli population will become immune to COVID-19 by the end of November, but the total number of death outcomes in Israel may approach 5,000.
- $R_0 = 1.10$ . This more severe scenario will increase the daily infection rate up to 9,500 cases in mid-October before starting a decrease down to 8,000 by the end of November due to an increasing percentage of population immunity. At its peak, it will require more than 2,000 respirators. About 16,000 people are expected to die from the virus during the simulated period, with at least 12% of the population getting immunity.
- $R_0 = 1.20$ . This is the most extreme and hopefully the least realistic scenario we have considered. Its peak is expected in August, when we may expect about 33,000 new cases per day requiring as many as 7,500 respirators, well above the current capacity of the Israeli hospitals. However, by the end of November, the

daily infection rate should go down to less than 1,000, due to at least one-third of the population becoming immune to the virus. This radical scenario may result in a death toll of about 40,000 people.



**Fig 7. Projected Daily Infection Rate in Israel.** Up to November 30, 2020



**Fig 8. Projected Daily Infection Rate in Israel.** Up to November 30, 2020

## Conclusion

The future evolution of COVID-19 in Israel and elsewhere is hardly predictable. Until a vaccine becomes available, which may take many months from now, the governments will have to strike the tough balance between health and economic issues. Obviously, the implementation of the "exit strategy" by the Israeli Government will require a close monitoring of the reproduction number. It is unclear at this stage how long  $R_0$  can be kept below the value of 0.80, which apparently is needed for meeting the current "exit strategy" criteria. To ensure the country's return to a proper level of economic activity, it may be necessary to tolerate higher values of  $R_0$  up to 1.0 and even 1.05, and we

would strongly suggest the Israeli Ministry of Health to make the necessary preparations for such a scenario. However, a further increase in  $R_0$  beyond the value of 1.05 should certainly be avoided by temporarily restoring the lockdown measures. 297  
298  
299

A further progress in understanding the current pandemic will be possible with a release of detailed clinical records of COVID-19 patients in Israel and other countries to the research community. Analyzing these records with state-of-the-art statistical and machine learning algorithms may reveal answers to many important epidemiological questions such as an accurate and early detection of high-risk patients, identification of the most infectious persons and locations, characterization of the most common infection pathways, etc. Many thousands of human lives worldwide are at risk and, as we all know, "data saves lives". 300  
301  
302  
303  
304  
305  
306  
307

## Acknowledgments 308

I thank the Israeli Ministry of Health, the Gertner Institute, and the Johns Hopkins University Center for Systems Science and Engineering for publishing the COVID-19 data used in this research. I am also grateful to Michal Feldman, Neil Gandal, Avraham Tabbach, Matan Yonas, Ady Pauzner, Sivan Frenkel (Tel-Aviv University), Yoram Louzoun (Bar-Ilan University), and Lev Muchnik (Hebrew University of Jerusalem) for the fruitful discussions of some of the topics presented in this paper. 309  
310  
311  
312  
313  
314

## References

1. Humanitarian Data Exchange (HDX)  
<https://data.humdata.org/event/covid-19>
2. COVID-19 Data Repository of the Israeli Ministry of Health  
<https://data.gov.il/dataset/covid-19>
3. Finkenstädt BF, Grenfell BT. Time series modelling of childhood diseases: a dynamical systems approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 2000;49(2):187–205.
4. Anderson RM, Anderson B, May RM. *Infectious diseases of humans: dynamics and control*. Oxford university press; 1992 Aug 27.
5. Maier BF, Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*. 2020 Apr 8.
6. Encyclopaedia Britannica. Case Fatality Rate — Definition, Example, Equation, Facts. <https://www.britannica.com/science/case-fatality-rate>
7. Kissler S, Tedijanto C, Goldstein E, Grad Y, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the post-pandemic period. *Science* 14 Apr 2020: eabb5793. DOI: 10.1126/science.abb5793.
8. Housen T, Parry AE, Sheel M How long are you infectious when you have coronavirus? *The Conversation* April 13, 2020 6.20am AEST.  
<https://theconversation.com/how-long-are-you-infectious-when-you-have-coronavirus-135295>
9. The Gertner Institute COVID-19 Scenarios - April 06, 2020

10. Modi C, Boehm V, Ferraro S, Stein G, Seljak U. Total COVID-19 Mortality in Italy: Excess Mortality and Age Dependence through Time-Series Analysis. medRxiv. 2020 April 20. DOI: <https://doi.org/10.1101/2020.04.15.20067074>.