Evaluation of South Korean NPI against COVID-19 by Effective Reproductive Number

Shin, Ikgyu Tom

It has been the best of times and it has been the worst of times. Best of times- cutting edge technologies, big data, the Fourth Industrial Revolution have arisen and shown plethora of opportunities and enlightening chances in future filled with happiness and the impossible to be the possible; that's what we've all thought of, at least for certain period of time. The worst of times, of course, came the outbreak of COVID-19.

COVID-19, a disease that causes respiratory illness, is derived from SARS-CoV-2 virus (Severe Acute Respiratory Syndrome Coronavirus 2), a member of Coronaviruses (Covs). SARS and MERS are also well-known family members of the family of coronaviruses but differ from COVID-19 in a sense that they caused serious upper respiratory tract illnesses resulting in high death rates. Low death rate of COVID-19 led the pandemic to last for almost 3 years already – vaccines are imperfect and there is no guarantee where the recovered individuals are fully protected from being reinfected.

Given such disastrous proliferation, countries all around the world initiated their own virus-protocol measures in order to prevent, protect, and contain the unknown virus from their people. Vaccine, perhaps, is one most efficient, effective means in such protection; however, with no information and data of any sort, it is nearly impossible for the manufacturers to fully supply and vaccinate the whole susceptible individuals. Here the concept of Non-Pharmaceutical Interventions comes in.

Non-Pharmaceutical Interventions, NPI in short, are actions individuals, communities, and even a nation could take in order to prevent and alleviate the spread of a pandemic besides the medical treatment including vaccines. Many interventions exist but here we will be majorly focusing on mitigation and suppression policies. Mitigation policy, in general, aims for the harmonization of the people along with the virus. Mitigation implements mild transmission reduction strategy in increasing awareness of personal hygiene and isolation of the exposed & infectious individuals from the susceptible. Simply put, it is targeting for gradual gain in herd immunity rather forcing radical policies in retaining influenza. On the other hand, suppression policy puts strong emphasis on containing virus' spread via individual tracking, frequent testing, restraining social and public activities until, ironically, supply of pharmaceutical interventions.

Suspicion and criticism reached its peak against European countries when they countries expressed their tendency in reinforcing herd immunity by implementing mitigation policies. Of course, whichever strategy would it be, leaning towards one sole policy with great bias would not efficiently solve the situation. In considering such logic, here in this paper, we will be focusing on South Korean policy where it implemented a combination of suppression and mitigation policy.

With such a dense population, zero to no knowledge of influenza, and failure in establishing protocol beforehand, the Korean government and the Ministry of Health and Welfare of South Korea faced their fate with a rough start. As a response, the authorities had regarded personnel's behavior to be the major contributing factor of the spread and as a result they have initiated phased social distancing policies. The social distancing policies were enacted and enforced in terms of the nation's first-level administrative divisions and were flexibly designed in which they could either be interpreted as suppressive or mitigative. Following is a simple chart visualizing South Korean NPI strategies.

Table 1. South Korean NPI policies per phase

Date	NPI	Policy Level	Phase	Notation	Group #	Closing Hour	ES & MS	HS	office	social event	miltary
03/22/2020	suppression	3	1	3_1	N/I	N/I	2/3 of all	2/3 of all	Flex	N/A	allowed
05/06/2020		3	1	3_1	N/I	N/I	2/3 of all	2/3 of all	Flex	N/A	allowed
08/19/2020		3	2	3_2	N/I	21	under 1/3	under 2/3	partial	in:100/out:50	prohibited
08/30/2020		3	2.5	3_2.5	N/I	21	under 1/3	under 2/3	partial	in:100/out:50	prohibited
09/14/2020		3	2	3_2	N/I	21	under 1/3	under 2/3	partial	in:100/out:50	prohibited
10/12/2020		3	1	3_1	N/I	Flex	under 1/3	under 2/3	Flex	100	allowed
11/19/2020		5	1.5	5_1.5	Flex	Flex	2/3 of all	2/3 of all	partial	100	allowed
11/24/2020		5	2	5_2	Flex	21	1/3 of all	2/3 of all	partial	100	allowed
12/08/2020		5	2.5	5_2.5	4	21	1/3 of all	1/3 of all	partial	50	prohibited
02/15/2021		5	2	5_2	4	21	1/3 of all	2/3 of all	partial	100	allowed
07/12/2021		4	4	4_4	4	21	2/3 of all	all	partial	50	allowed
11/01/2021	mitigation	N/A	N/A	N/A	4	21	N/A	N/A	N/A	N/A	N/A
12/18/2021		N/A	N/A	N/A	6	22	N/A	N/A	N/A	N/A	N/A
02/19/2022		N/A	N/A	N/A	6	23	N/A	N/A	N/A	N/A	N/A
03/05/2022		N/A	N/A	N/A	8	23	N/A	N/A	N/A	N/A	N/A
03/20/2022		N/A	N/A	N/A	10	24	N/A	N/A	N/A	N/A	N/A

N/A = No particular limitations

[military] soldier vacation allowed | fully allowed or partially allowed

N/I = No exact information

[ES & MS, HS] Elementary School | Middle School | High School

Flex = Flexible max capacity in terms of area

[social event] wedding, funeral, etc.

in:#/out:# Max capacity for either indoor or outdoor events

[Group #] Max capacity for gathering

[Notation] Notation used for **table 2 & 3**

The [Date] column from figure 1 is the starting date of the corresponding phase of social distancing policy which is in the second column. [Closing Hour] stands for the closing time for businesses depending on their business type. [Policy Level] (temporarily named) stands for the social distancing policies' distribution of phases; the phases had been divided into 3, 4, and 5 phases depending on the trend.

To briefly introduce the structure of the paper, R_0 will be calculated and visualized through the next generation method; SEIR model will be used in representing COVID-19 which will also be discussed further in latter portion. Methodology in assessing efficiency of South Korean NPI will be analysis of the R_t trend with time as an X-axis, strong emphasis on a date near [Policy Level] adjustments. As for mitigation phases, we will compare the death rate from both pre- and post-COVID era; since South Korean government's mitigation policy puts emphasis on securing severely affected individuals as a priority, where suppression majorly aimed for perfectly inhibiting the spread, we have decided death rate to be proper variable to consider.

R_0, the basic reproductive number, is a great measure in calculating the potential transmission or a spread of a disease, commonly known as the average of number of secondary infections produced by an infection in a population (where everyone is susceptible.) In deriving basic reproduction number R_0, I have implemented Next Generation Method (NGM) [Equation 1.5~1.8], which is useful in deriving R_0 mostly from compartmental models where population can be allocated into separate categories. I have chosen Susceptible-Exposed-Infectious-Recovered (SEIR) [Equation 1.1~1.4] model for COVID-19. Assumptions are $S+E+I+R=N$ where N is a total population size being constant, homogeneity, and the fact that recovered individuals from R compartment are permanently immune.

$$\frac{dS}{dt} = -\beta SI \quad (\mathbf{1.1})$$
$$\frac{dE}{dt} = \beta SI - \sigma E \quad (\mathbf{1.2})$$
$$\frac{dI}{dt} = \sigma E - \gamma I \quad (\mathbf{1.3})$$
$$\frac{dR}{dt} = \gamma I \quad (\mathbf{1.4})$$

$$F = \begin{pmatrix} \frac{d\beta SI}{dE} & \frac{d\beta SI}{dI} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \beta S \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \beta \\ 0 & 0 \end{pmatrix} \quad (\mathbf{1.5})$$

$$V = \begin{pmatrix} \frac{d\sigma E}{dE} & \frac{d\sigma E}{dI} \\ \frac{d\gamma E - \sigma E}{dE} & \frac{d\gamma E - \sigma E}{dI} \end{pmatrix} = \begin{pmatrix} \sigma & 0 \\ -\sigma & \gamma \end{pmatrix} \quad (\mathbf{1.6})$$

$$F \cdot V^{-1} = \begin{pmatrix} 0 & \beta \\ 0 & 0 \end{pmatrix} \cdot \frac{1}{\sigma\gamma} \begin{pmatrix} \gamma & 0 \\ \sigma & \sigma \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\gamma} & \frac{\beta}{\sigma} \\ 0 & 0 \end{pmatrix} \quad (\mathbf{1.7})$$

$$|F \cdot V^{-1}| = R_0 = \frac{\beta}{\gamma} \quad (\mathbf{1.8})$$

Derivation results in $R_0 = \frac{\beta}{\gamma}$ where β and γ stands for transmission rate and recovery rate, respectively. However, since both parameters are not easily computable, I have decided to implement another method in deriving exact value for R_0 with the South Korean COVID-19 test positive population dataset. I have utilized Lotka's continuous equation [Equation 2.1], the study of age-structured population growth, and substituted appropriate variables to derive desired R_0 value.

$$1 = \int_0^\infty e^{-ra} n(a) da \quad (\mathbf{2.1})$$

$$R_0 = \int_0^\infty n(a) da \xrightarrow{yields} n(a) = R_0 \times g(a) \quad (\mathbf{2.2})$$

While $n(a)$ from Equation (2.1) stands for birth rate of age group a, Jeong replaces its definition with total number of patients an infected individual can infect during the infectious phase (Jeong, 2020).

$$\frac{1}{R_0} = \int_{a=0}^\infty e^{-ra} g(a) da \quad (\mathbf{2.3})$$

After substitution, the right side of Equation 2.3 resembles moment generating function with random variable T and probability density function g which can now be represented as Equation (2.4).

$$M_T(z) = E(e^{tT}) = \int_{a=0}^{\infty} e^{-ra} g(a) da \quad \mathbf{(2.4)}$$

$$\frac{1}{R_0} = M_T(-r) \xrightarrow{yields} R_0 = \frac{1}{M_T(-r)} \xrightarrow{yields} \widehat{R_0} = \frac{1}{M_T(-\hat{r})} \quad \mathbf{(2.5)}$$

That is, if the estimated value for the rate of increase r and the probability density function of the distribution of generation time T are known, $\widehat{R_0}$ will be an estimate for R_0[Equation 2.5].

$$\frac{d_i(t)}{dt} \propto i(t) \xrightarrow{yields} \frac{d_i(t)}{dt} = ri(t) \xrightarrow{yields} i(t) = i_0 \exp(rt) \quad \mathbf{(2.6)}$$

$$\hat{r} = \arg min_r \left[\sum_{t=0}^{T} |i_t^{Data} - i(t)|^2 \right] = \arg min_r \left[\sum_{t=0}^{T} |i_t^{Data} - i_0 \exp{(rt)}|^2 \right] \quad \mathbf{(2.7)}$$

Here, Jeong incorporates Malthusian model, an exponential model for projecting population growth, and derives deterministic model [Equation 2.6] where $i(t)$ is differentiable function for the number of incidences at time t along with proportional constant r and $i(t = 0)$ being i_0. Then right end side equation from Equation (2.6) can be represented as Equation (2.7), a method of least squares for time series data $i_0^{Data}, \dots\ i_t^{Data}$.

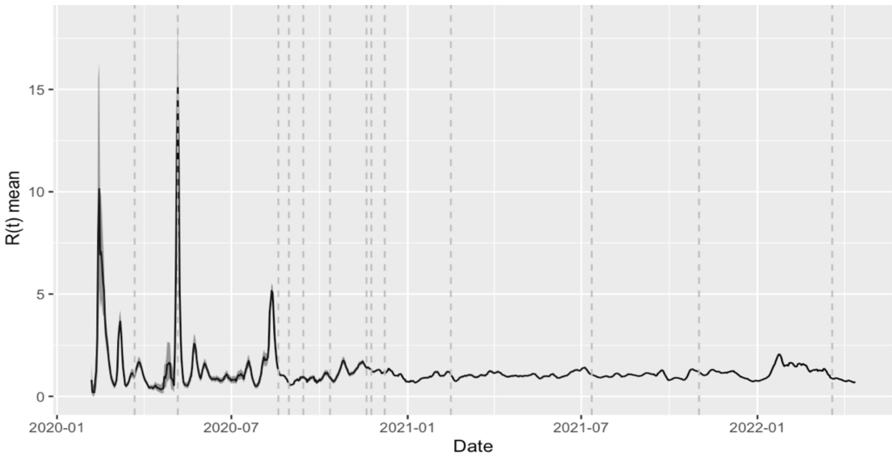
While Jeong has applied serial interval of 4.5(1.4, 7.6) with 0.69 as coefficient of variation, assuming Gamma distribution, I have decided to use 5.19(3.8, 6.58) and 0.2678 for coefficient of variation which are pooled estimates from random effect model on meta-analysis on many existing literatures, mostly analysis on East Asia region; due to existing heterogeneity from the studies, random effect model is assumed to be more appropriate (Rai et al, 2020). I have applied three scenarios by setting three different dates for the outbreak (2020/02/14, 2020/02/15, and 2020/02/16); though the first confirmed case arose on 2020/01/21, I found full-scale spread originated from the super spreader, a member of Shincheonji Church of Jesus who tested positive around 2020/02/16.

$$\widehat{R_0} = (1 + 5.19 \cdot \hat{r} \cdot 0.2678^2)^{1/0.2678^2} \quad \mathbf{(2.8)}$$

In obtaining $\hat{r}$, I have applied bootstrapping on dataset (filtered South Korea from global time series data of confirmed cases crawled from Johns Hopkins Github) with 1,000 replications, assuming Poisson distribution with mean i_t^{Data} via R program. $\widehat{R_0}$ was then calculated by Equation (2.8). As a result, I obtained 0.27, 0.28, 0.29 for $\hat{r}$ and 3.79, 3.99, and 4.23 for $\widehat{R_0}$, respective to each scenario dates. Results seem to be higher than the literature possibly due to different starting period and different serial interval; I found the results to be plausible in a sense that they successfully depict explosive proliferation in earlier dates.

In assessing Korean NPI against COVID-19 I computed daily trend R_t and compared those values in between changes in social distancing policy. R_t, the effective reproductive number is the measure which depicts the number of individuals who become infected per infectious person at time t; here, recognize the basic reproductive number, R_0, can be noted as R_t when t=0. First, I have implemented estimate_Rt() function from sars2pack for simple overview of a trend with the same dataset but with the whole time period in between the outbreak and today.

Figure 2. R_t trend divided by South Korean NPI phases [estimate_Rt()]



Data from : [https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv]

Results from estimate_Rt() were difficult to interpret due to the lack of documentation regarding package's design or method; however, we can see overall trend how R_t fluctuates in the beginning phase of pandemic. We can also see how shifts in NPI policy levels strongly restrain explosive spread by simply looking at the trend of the black line in between grey dotted line [Figure 1]. We can also see how shifts in NPI policy levels strongly restrain explosive spread then another strong pull down around days with resurging trend

Due to lack of confidence in the first attempt's result, I have also implemented a method created by Kevin Systrom based on a paper by Bettencourt & Riberio.

$$P(R_t|k) = \frac{P(k|R_t) \cdot P(R_t)}{P(k)} \quad (3.1)$$

$$P(R_1|k_1) \propto P(R_1) \cdot \mathcal{L}(R_1|k_1) \quad (3.2)$$

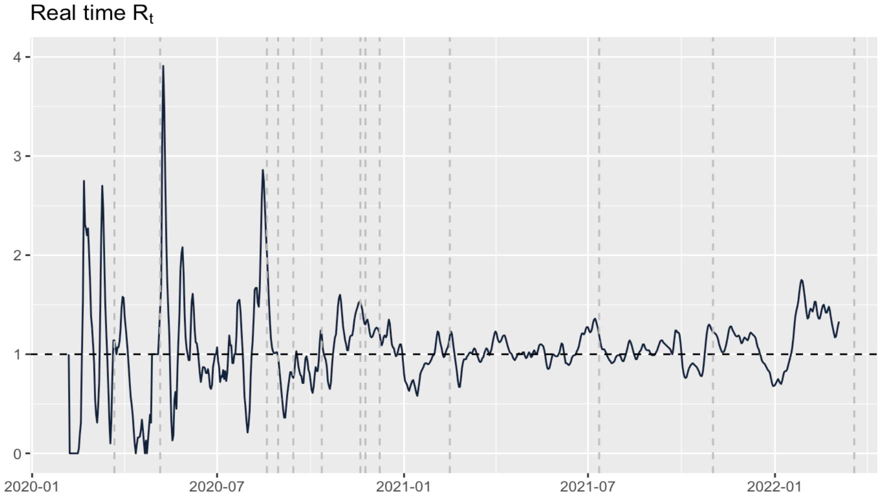
$$P(R_2|k_1, k_2) \propto P(R_2) \cdot \mathcal{L}(R_2|k_2) = \sum_{R_1} P(R_1|k_1) \cdot P(R_2|k_1) \cdot \mathcal{L}(R_2|k_2) \quad (3.3)$$

$$P(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \xrightarrow{yields} \lambda = k_{t-1} e^{\gamma(R_t - 1)} \xrightarrow{yields} P(k|R_t) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (3.4)$$

The methodology follows Bayes' rules under theory that value of R_t depends on the value of R_{t-1}(Bettencourt & Riberio, 2008). Equation (3.1) is a likelihood of observing k new cases given R_t, multiplied by theoretical prior value of $P(R_t)$ divided by probability of general positive cases' observations. Equations (3.2 & 3.3) are simply first two days within numerous iterations, assuming Gaussian distribution of R_t centered around R_{t-1} with $\mathrm{N}(R_{t-1}, \sigma)$: here σ is a hyperparameter which needs to be a value that maximizes $P(k)$. Then Bettencourt & Riberio brings in Poisson distribution with λ new cases per day and k being probability in observing new cases then links λ to R_t ending up with reformation of likelihood function with Poisson distribution by fixed k and varying R_t: here, γ is a reciprocal of serial interval, 5.19 [Equation 3.4]. Computation has been done via R program.

Results from shadowing Bettencourt & Riberio were more credible with no eye catching variances; we can also see overall trend, similar to that of estimate_Rt() where R_t fluctuates in the beginning phase of pandemic. Again, we can also see how shifts in NPI policy levels strongly restrain explosive spread then another strong pull down around days with resurging trend [Figure 2].

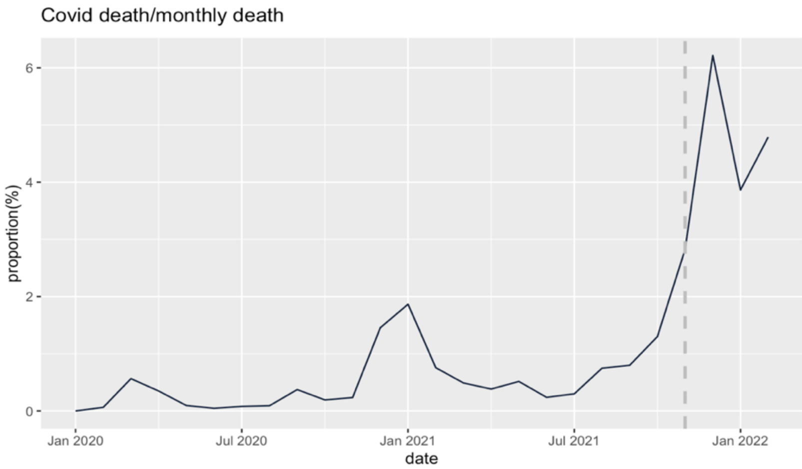
Figure 2. R_t trend divided by South Korean NPI phases [Bettencout & Riberio]



Data from : [https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv]

While suppression method can be easily assessed by exhibiting R_t trend by NPI phase, since mitigation method has no intention in suppressing infected individuals, rather protecting death, I have computed total death population and death population from COVID-19. I simply divided number of dead individuals from COVID-19 by total death population and multiplied it by 10 since the values were too minute to compute. Interestingly, we can see how proportion of COVID-19 death rises in recent dates which might represent failure in South Korean mitigation method [Figure 3].

Figure 3. Death trend comparison between Suppression and Mitigation



Data from : [KDX & https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_global.csv]

Though the paper might wrong computational assumptions and improper methodology, we are able to see a $R_t$trend how South Korean NPI was effective in suppressing positive incidences until vaccination yet quite ineffective in mitigating the pandemic; however, having R_t value around 1.1, which is a huge decrement compared to that of the beginning phase, we could positively expect herd immunization. As for future edit, I could study more of assumptions and statistical distribution to yield more precise result. By then, I would expect well cleaned data and clear parametric values.

References

Bettencourt LMA, Ribeiro RM (2008) Real Time Bayesian Estimation of the Epidemic Potential of Emerging Infectious Diseases. PLoS ONE 3(5): e2185. https://doi.org/10.1371/journal.pone.0002185

Chowell G. 2017. Fitting Dynamic Models to Epidemic Outbreaks with Quantified Uncertainty: A Primer for Parameter Uncertainty, Identifiability, and Forecasts. Inf Dis Model. doi:10.1016/j.idm.2017.08.001.

Jeong, J., Kwon, H., Hong, S., Lee, M. 2020. Estimation of Reproduction Number for COVID-19 in Korea https://doi.org/10.7469/JKSQM.2020.48.3.493

Rai, B., Shukla, A., & Dwivedi, L. K. (2021). Estimates of serial interval for COVID-19: A systematic review and meta-analysis. *Clinical epidemiology and global health*, *9*, 157–161. https://doi.org/10.1016/j.cegh.2020.08.007

Systrom K. (2020) [Estimating COVID-19's Rt in Real-Time]