Understanding the Time Dependent Transmissibility of COVID-19 Through Its Effective Reproduction Number \((R_t)\)

See Jie, Rajasvaran Logeswaran

Asia Pacific University of Technology and Innovation, 57000 Kuala Lumpur, Malaysia.

ABSTRACT

Background: Over six centuries ago, the most fatal pandemic ever recorded in human history, the Black Death (The Plague), emerged with an estimated death toll of 75-200 million. At present, COVID-19 has surfaced causing over 5 million infected and 340 thousand deaths worldwide, with the numbers still rising.

Objective: This paper focuses on the significance of the effective reproduction number \((R_t)\) in the epidemiology of SARS-CoV-2, to understand the time-dependent transmission pattern of this virus after intervention measures such as city lockdowns and social distancing. Studies done by various parties on estimating the \(R_t\) for different countries are discussed and interpreted.

Results: It is found that all the countries studied still show an \(R_t\) of greater than 1, indicating that the COVID-19 outbreak is still ongoing. However, some countries are gradually keeping COVID-19 under control with a decreasing \(R_t\) after implementing intervention measures.

Conclusion: COVID-19 is still spreading rapidly across the entire globe, and the soonest production of vaccines is expected to be early 2021, but it is estimated that it will take up to five years to be developed.

Key Words: COVID-19, SAR-Cov-2, Effective Reproduction Number, Time-Dependent Transmissibility

INTRODUCTION

It was not long ago when the World Health Organization (WHO) declared the Coronavirus Diseases 2019 (COVID-19) outbreak a pandemic, and as of 23rd May 2020, there have been 5,346,876 confirmed cases and 340,869 deaths worldwide, covering 213 countries and territories.\(^1\) Firstly emerged as an epidemic on the 31st of December 2019 in Wuhan, Hubei Province in China, several Chinese scientists isolated and sequenced the genome of the virus.\(^2\) The virus was tentatively named by the WHO as the 2019 novel coronavirus (2019-nCoV) and later identified as the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), causing the Coronavirus Disease 2019 (COVID-19). According to the COVID-19 can affect certain groups which include the older populations and individuals who suffer from underlying health conditions that COVID-19 hypothetically induces mild symptoms to other respiratory infections that exhibited an ability to generate severe diseases towards the affected group.\(^3\)

It was further explained that a clear understanding of the epidemiology of the virus was still being elucidated and the estimation of basic reproduction number \((R_0)\) as well as effective reproduction number \((R_t)\) was vital.

Basic Reproduction Number \((R_0)\)

In epidemiology, it is essential to estimate the actual or possible outbreaks of a particular disease, such as the Human Immunodeficiency Virus (HIV),\(^4\) Severe Acute Respiratory Syndrome (SARS),\(^5\) Tuberculosis (TB),\(^6\) and the recent Coronavirus Disease 2019 (COVID-19). The estimation of the possible outbreak during the start of an epidemic is usually represented by a parameter called Basic Reproduction Number \((R_0)\), which is widely used in assessing the transmissibility of pathogens as well as estimating the severity of the outbreak. \(R_0\) is defined as the expected number of secondary cases directly generated by one case in a population, wherever all individuals are susceptible to infection without any deliberate intervention in disease transmission.\(^7\) \(R_0 > 1\) indicates the outbreak of a disease, whereas \(R_0 < 1\) means that...
the pathogen will eventually disappear. Although $R_t$ does not explain the speed of spread after intervention controls such as social distancing and vaccination have been implemented, it is directly related to the type and intensity of interventions needed during the early stage of an outbreak. Upon successful development of the vaccine, $R_0$ can be used to determine the critical coverage of immunization required to eradicate the disease. The critical vaccination coverage ($\rho_c$) can be estimated through Eq. (1), where $\varepsilon$ is the direct effect of vaccination.

$$\rho_c > \frac{1}{\varepsilon} \left( 1 - \frac{1}{R_0} \right)$$  \hspace{1cm} (1)

$\rho_c$ suggests the proportion of susceptible individuals be vaccinated, and the protection conferred to the population achieving the coverage is known as herd community.\(^8\)

**Effective Reproduction Number ($R_t$)**

Although $R_0$ is essential to understand the transmissibility of a disease and interventions strategies needed, it is estimated based on assumptions that the population is fully susceptible, which only happens during the beginning of an outbreak. During an outbreak, it is important to evaluate the time-dependent transmission potential of an infectious disease. This is estimated by calculating the Effective Reproduction Number ($R_t$), i.e. the actual average number of secondary cases per primary case at time $t$, for $t > 0$.\(^9\) $R_t$ is commonly used to characterize pathogen transmissibility during an epidemic due to its time and situation specificity,\(^1\) that takes into account the decline in susceptible individuals (intrinsic factor) and control measures implementation (extrinsic factor).\(^12\) $R_t < 1$ suggests that the disease outbreak has been brought under control at time $t$, whereas $R_t > 1$ indicates that the outbreak is still ongoing. To bring an epidemic under control, the monitoring of $R_t$ overtime provides feedback on the effectiveness of interventions and the importance to further intensify control efforts.\(^13,14\) Given that the goal of intervention effort is to bring $R_t$ below the threshold value of 1 and subsequently close to 0 as soon as possible.

**Estimation of $R_t$ for COVID-19**

The study\(^15\) carried out $R_0$ and $R_t$ estimations in South Korea, but the cases are split into: total, Sincheonji, except Sincheonji, Daegu-Gyeonbuk and except-Daegu-Gyeonbuk subgroups, to thoroughly understand the transmission patterns of major cluster cases in South Korea. The exponential growth model\(^16\) was first fitted on the $R_0$ values, followed by sensitivity analysis done on the exponential growth period. Due to the unknown generation time (GT) of COVID-19 at the time of the study, the serial interval was modelled using gamma distribution with mean ± standard deviation of $2.0 \pm 1.0$, $3.0 \pm 1.5$, $4.0 \pm 2.0$, $5.0 \pm 2.5$, and $6.0 \pm 3.0$ days, obtained by referring to the incubation period of two to 14 days.

An initial $R_t$ for total cases was estimated to be around 3.4. However, the outcome of the sensitivity test suggests that the $R_0$ value is highly sensitive to the GT of COVID-19, hence the estimated $R_t$ might not be reliable as the GT used was just based on assumptions. After the Sincheonji mass infection incident, the health authorities of Korea implemented a series of intervention procedures, including active COVID-19 testing, isolation, social distancing, school closures and active surveillance. By applying the same method, the $R_t$ was estimated and the result is illustrated in Figure 1. Although the $R_t$ of total cases show a downward trend, both “Except Sincheonji” and “Except Daegu-Gyeonbuk” clusters were still fluctuating, suggesting the possibility of another mass infection if not dealt with carefully. Nonetheless, thanks to the intervention approach, $R_t$ of most of the cluster groups gradually decreased. However, the estimated $R_t$ may be different from the actual $R_t$ due to the unknown GT, Figure 1.

![Figure 1: $R_t$ estimation of different subgroups\(^15\).](image)

The exponential growth rate (EG) and time-dependent (TD) methods were employed using local transmission cases observed in Germany, France, Italy and Spain from 20th February 2020 to 9th March 2020.\(^17\) Under the EG method, the exponential growth stage of the outbreak is first fitted using a Poisson regression model, under the assumption that the GT of COVID-19 follows a gamma distribution.\(^18\) The $R_t$ is then calculated through the transformation of the EG rate. Through the TD method, $R_t$ is computed by averaging the overall transmission networks that are compatible with the observed epidemic curve.\(^18\) The TD method takes into account the yet-unrecorded cases because it uses a Bayesian statistical framework. Both EG and TD methods are highly dependent on the value of the GT, measured by the onset time lag between primary and secondary causes. As this is difficult to obtain, GT is assumed to be equal to the incubation period of 5.8 days with a standard deviation of 2.6 days.\(^20\) Sensitivity analysis is then carried out using an in-
cubation period of 4 days with a standard deviation of 2.4 days.\textsuperscript{21} From the results in Table 1 for the different models up until 9\textsuperscript{th} March 2020, the \( R_t \) in all four countries were greater than two, indicating the progression of the COVID-19 outbreak. Stricter control measures are needed in these countries, especially France, Germany and Spain, to bring the pandemic under control. Besides that, lockdown measures had just been implemented in Italy and therefore \( R_t \) should be constantly monitored to understand the effectiveness of the measure.

Some limitations of this study include that the models are highly dependent on the GT of COVID-19, which is relatively difficult to estimate. Further, the reported cases may be lower than the actual cases, which can affect the accuracy of the study. Despite the limitations, this study provides important findings of the transmissibility of COVID-19 in the four countries. The European Centre for Disease Prevention and Control (ECDC) should work closely with these countries, Table 1.

<table>
<thead>
<tr>
<th>Country</th>
<th>Start Date</th>
<th>End Date</th>
<th>( R_t ), GT = 5.6 days</th>
<th>( R_t ), GT = 4 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>02/23/2020</td>
<td>03/09/2020</td>
<td>3.27</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(3.17, 3.38)</td>
<td>(2.25, 2.35)</td>
</tr>
<tr>
<td>France</td>
<td>02/23/2020</td>
<td>03/09/2020</td>
<td>6.32</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(5.72, 6.99)</td>
<td>(3.37, 3.85)</td>
</tr>
<tr>
<td>Germany</td>
<td>02/21/2020</td>
<td>03/09/2020</td>
<td>6.07</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(5.51, 6.69)</td>
<td>(3.28, 3.74)</td>
</tr>
<tr>
<td>Spain</td>
<td>02/19/2020</td>
<td>03/09/2020</td>
<td>5.08</td>
<td>3.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4.51, 4.74)</td>
<td>(2.86, 3.37)</td>
</tr>
</tbody>
</table>

In the \( R_t \) of COVID-19 for 32 countries outside of China with over 100 reported cases by 13\textsuperscript{th} March 2020 were computed by applying the EG method \textsuperscript{[16]}. Data on the daily reported new COVID-19 cases in these countries were used, alongside with assumption of mean serial interval of 4.7 days and standard deviation of 2.9 days.\textsuperscript{22, 23} Of these 32 countries, United States (2,294), France (3,671), Germany (3,675), Spain (5,232), Korea (8,086), Iran (11,364) and Italy (17,660) had more than 1,000 reported COVID-19 cases. Bahrain, Slovenia, Qatar, Spain, Denmark and Finland had \( R_t \) greater than four. The estimated \( R_t \) of all 32 countries as of 13\textsuperscript{th} March 2020 are shown in Table 2. The \( R_t \) should be constantly monitored like some of the countries with relatively low \( R_t \) (<4) have been reported to have a drastic increase of COVID-19 cases in the span of two months (from March to May). By 13\textsuperscript{th} May 2020, the country with the highest number of reported cases (1,408,636) in the world was the United States, with a \( R_t \) of 3.29. One of the reasons for the significant increase might be the lack of awareness of both the government and citizens, Table 2.

By using the daily reported cases, as released by the Korea Centers for Disease Control and Prevention (KCDC) including locally transmitted and imported cases of COVID-19 in South Korea from 20\textsuperscript{th} January to 26\textsuperscript{th} February 2020.\textsuperscript{24} estimated the \( R_t \) in Korea. To estimate the growth rate of the pandemic, it is crucial to identify the epidemic curve according to the dates of symptoms onset rather than the dates of reporting. However, the symptoms onset dates were only available for the first 28 imported cases. To handle this problem, symptoms onset dates were imputed by constructing 300 epidemic curves by dates of symptoms onset and obtaining a mean incidence curve after dropping the last three data points to adjust for reporting delays. By using the generalized growth model (GGM) to characterize the incidence curves for local and imported cases.\textsuperscript{25} This
model describes the growth profile with the help of growth rate parameter \((r)\) and scaling of growth rate parameter \((p)\), where the GT “is assumed to follow a gamma distribution with a mean of 4.41 days and standard deviation of 3.17 days.\(^{25,26}\) The discretized probability distribution \((\rho_i)\) of the generation interval is then simulated from the GCM and applied on:

\[
R_t = \frac{I_i}{\sum_{j=0}^{\infty} \frac{J_j}{(t_j+j+\alpha J_j)} \rho_i}
\]

(2)

where, \(I_i\) is the local incidence at the time \(t\), \(J_j\) is the local incidence at the time \(t\), and \(1<\alpha<1\) quantifies the relative contribution of imported cases to secondary disease transmission.\(^{15}\) Finally, the \(R_t\) is estimated “for 300 simulated curves assuming a Poisson error structure to derive the uncertainty bounds around the \(R_t\) curve”.\(^{27}\) The \(\hat{R}_t\) was estimated to be 1.5 with 95% Confidence Interval (CI) of (1.4, 1.6) as of 26\(^{th}\) February 2020, and the estimates indicate a sustained transmission of COVID-19 in Korea. This finding is lower compared to the \(R_t\) estimated,\(^{15}\) which gave a \(R_t\) for total cases of more than 2 as of 26\(^{th}\) February 2020.

The impact of the intervention on the spread of COVID-19 in India was studied in \(^{28}\) by applying a compartmental model that separated the total human population \((N)\) into susceptible \((S)\), exposed \((E)\), symptomatic \((I)\), asymptomatic \((I_a)\), quarantined \((I_q)\), hospitalized \((H)\) and recovered \((R)\), hence:

\[
N = S + E + I + I_a + I_q + H + R.
\]

(3)

After close contacts with symptomatic or asymptomatic individuals, the susceptible population will become exposed. It is assumed that the symptomatic individuals have a higher transmission rate than asymptomatic individuals and the rate of a new infection is given by Eq. (4), where \(\beta\) is the transmission rate of the disease and \(n(<1)\) is the modification parameter for the reduction in the transmission rate from asymptomatic individuals:

\[
Rate\ of\ new\ infection = \frac{\beta S I (I+n I_a)}{N}
\]

(4)

Of all exposed individuals, \(p_1\) the fraction will be symptomatic, \(p_2\) fraction will be asymptomatic, whereas \(p_3 = 1 - p_1 - p_2\) will be quarantined. Symptomatic individuals will show severe symptoms after \(\alpha^{-1}\) time and will be hospitalized. On the other hand, asymptomatic individuals will naturally recover at a rate of \(r_a\). Quarantined individuals will either naturally recover at a rate of \(r_q\) or become infected and hospitalized at a rate of \(r_q\). Finally, hospitalized individuals will either recover at a rate \(r\), or die at a rate of \(d\). The full transmission dynamics of this COVID-19 disease is illustrated in Figure 2.

The daily new infection curve of reported COVID-19 cases is then used to estimate \(\hat{R}_t\) by applying the following formula, which is derived from the renewal equation of the birth process:

\[
R(t) = \frac{b(t)}{\int_0^\infty b(t-\alpha)g(\alpha)d\alpha}
\]

(5)

where, \(b(t)\) is the number of new cases on day \(t\), and \(g(\alpha)\) is the generation interval distribution for COVID-19. However, one drawback of the model is that it only assumes asymptomatic individuals to recover naturally and have neglected the fact that there are possibilities for asymptomatic individuals to start showing symptoms and will be quarantined, hospitalized or die from the disease. This is supported by\(^{29}\), were out of the 13 residents of the Long Term Care Nursing Facility, Washington who were asymptotically infected, 10 started developing symptoms after seven days.\(^{29}\)

The estimated \(\hat{R}_t\) of COVID-19 outbreak in India from 2\(^{nd}\) March to 24\(^{th}\) March 2020 is shown in Figure 3. Due to zero newly reported cases on 6\(^{th}\) March 2020, \(\hat{R}_t\) of zero is observed. A high \(\hat{R}_t\) of 7.54 and 8.86 are also observed on the 6\(^{th}\) and 17\(^{th}\) days due to the exceptionally high reported cases those two days. Further projected the estimated COVID-19 cases in India until September by taking into account intervention methods such as lockdown, awareness programs, proper hand sanitization, etc.\(^{28}\) It is observed that the stronger the intervention, the sooner the pandemic will be brought under control, Figure 3.
A semi-mechanistic Bayesian hierarchical model was used in \(^{30}\) to infer the impact of interventions across 11 countries in Europe, under the assumption that the changes in \(R_t\) are an immediate response to the effectiveness of interventions. To take into account the time lag between infection and death, the model estimates the changes in \(R_t\) by computing backwards from the deaths observed over time to estimate the transmissions that happened a few weeks prior. Besides, it is assumed that each type of intervention method will have the same effect on \(R_t\) overall countries, indicating that the results are strongly affected by the timing of intervention. In the infection model, the true number of infected individuals, \(c\), is modelled using the widely used discrete renewal process.\(^{31, 32}\) The serial interval distribution was modelled using gamma distribution of \(g\sim\text{Gamma}(6.5, 0.62)\), which is the time between when a person gets infected and the person infecting a second person.\(^{30, 31}\) The serial interval is assumed to be the same for all countries. A piecewise constant function that scales \(R_{t,m}\) from a baseline prior \(R_{0,m}\) and is driven by six different interventions, is \(k\), detailed as in Table 3.

<table>
<thead>
<tr>
<th>(k) value</th>
<th>Description of Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>school and university closures</td>
</tr>
<tr>
<td>2</td>
<td>self-isolating if ill</td>
</tr>
<tr>
<td>3</td>
<td>the banning of public events</td>
</tr>
<tr>
<td>4</td>
<td>any government intervention in place</td>
</tr>
<tr>
<td>5</td>
<td>implementing a partial or complete lockdown</td>
</tr>
<tr>
<td>6</td>
<td>encouraging social distancing and isolation</td>
</tr>
</tbody>
</table>

The indicator variable for intervention is denoted by \(I_{k,m,t}\), which will be one if intervention \(k\) has been implemented in country \(m\) at time \(t\), and 0 otherwise. The \(R_t\) in-country \(m\) at time \(t\) is then computed by Eq. (6), where \(\alpha\) is a parameter used to characterize \(R(t,m)\) with a chosen independent prior distributions of \(\alpha\sim\text{Gamma}(0.5, 1)\):

\[
R_{t,m} = R_{0,m} \exp - \sum_{k=1}^{6} \alpha_k I_{k,m,t}
\]  

(6)

The \(R_{0,m}\) is chosen to be \(R_0 \sim \text{Normal}(2.4, |k|)\), where \(k \sim \text{Normal}(0.0.5)\). Computed \(R_t\) for all 11 countries is then plotted, alongside the timeline of different intervention measures, giving a clear visualization of the impact of intervention methods. An example of the impact of intervention on \(R_t\) in Austria is illustrated in Figure 4. After undergoing sensitivity analysis, it is concluded that the choices of prior distribution do not affect the results from the main analysis. From Figure 4, it is apparent that \(R_t\) in Austria drops each time an intervention measure is implemented.\(^{32}\) On 16th March 2020, the \(R_t\) dropped drastically due to the implementation of 3 interventions on the same day, Figure 4.

![Figure 4: Impact of Intervention on \(R_t\) in Austria\(^{30}\).](image)

**CONCLUSIONS**

COVID-19 is still spreading rapidly across the entire globe, and the soonest production of vaccines is expected to be early 2021, but it is estimated that it will take up to five years to be developed. It is important to continue monitoring the \(R_t\) for different countries to completely understand the effect of the interventions and the potential need for other control measures. Based on the studies done in different countries to estimate \(R_t\), there is no doubt that intervention measures such as social distancing can be carried out and is essential to keep the outbreak under control. Apart from that, by estimating the time to time \(R_t\) countries that have successfully brought the spread of COVID-19 under control will be able to identify what interventions are the most effective, which can be used as a benchmark for other countries that are currently battling the disease. Once the vaccines are developed, it can also be identified by how fast the vaccine brings \(R_t\) to below one.

**ACKNOWLEDGEMENTS**

The authors also wish to express gratitude to the management of Asia Pacific University of Technology & Innovation (APU) for their support.

**Conflict of Interest**

The authors involved in the current study does not declare any competing conflict of interest.

**Funding and Sponsorship**

No fund or sponsorship in any form was obtained from any organization for carrying out this research work.
REFERENCES


