A simple model for fitting mild, severe, and known cases during an epidemic with an application to the current SARS-CoV-2 pandemic

Matthew I. Betti a, *, Jane M. Heffernan b

a Mathematics and Computer Science, Mount Allison University, Sackville, New Brunswick, Canada
b Mathematics and Statistics, Centre for Disease Modelling, York University, Toronto, Ontario, Canada

ARTICLE INFO

Article history:
Received 29 May 2020
Accepted 1 January 2021
Available online 15 January 2021
Handling editor: Dr. J Wu

Keywords:
Basic reproduction number
Epidemic curve fitting
COVID-19
Other words

ABSTRACT

One of the major difficulties with modelling an ongoing epidemic is that often data is limited or incomplete, making it hard to estimate key epidemic parameters and outcomes (e.g. attack rate, peak time, reporting rate, reproduction number). In the current study, we present a model for data-fitting limited infection case data which provides estimates for important epidemiological parameters and outcomes. The model can also provide reasonable short-term (one month) projections. We apply the model to the current and ongoing COVID-19 outbreak in Canada both at the national and provincial/territorial level.

1. Introduction

During a pandemic/epidemic, governments and media report on the total number of known cases to the public. These reports are done at an almost regular basis - during the most recent COVID-19 epidemic, daily reports are released. Using the reported data, modellers are often tasked to estimate important disease parameters, such as the basic reproduction number, attack rate, actual case load, and the effective reproduction number as mitigating strategies are implemented. This, however, is a difficult task since only a fraction of the total case-load is known, and these important disease parameters also depend on asymptomatic and/or mildly symptomatic non-observed infections.

The IDEA model (Fisman et al., 2013) has been shown to be effective in determining various epidemic parameters when data is limited. This model uses the basic reproduction number, \( R_0 \) to estimate the attack rate, epidemic duration, and turning points of an epidemic. The model has been used most notably to predict the aforementioned parameters for a influenza A pandemic (Fisman et al., 2013) and the 2014–2015 Ebola outbreak (Tuite & Fisman, 2018).

Other fitting methods and models that have been employed in past epidemics, and during the current COVID-19 outbreak, include reproduction number fitting through an SEIR model (Fang et al., 2020) or a delay SIR model (Anastassopoulou et al., 2020). Briefly, most methods apply statistical modeling and Bayesian inference to parameterize assumed epidemiological

* Corresponding author.
E-mail address: mbetti@mta.ca (M.I. Betti).
Peer review under responsibility of KeAi Communications Co., Ltd.

https://doi.org/10.1016/j.idm.2021.01.002
2468-0427/© 2021 The Authors. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
model structures While many of these methods and models produce accurate results, many lack the ability to differentiate between different types of infection, or the discrepancy between known cases and total cases. The former is important for informing policy on healthcare resources as more severe cases create a larger burden on the healthcare system. The latter discrepancy is an important aspect of epidemic fitting to address, due to the fact that unknown cases can still transmit infection, creating for discrepancies in estimation of the effective reproduction number.

In this study, we propose a simple ordinary differential equation model which can be used to track the dynamics of mild, severe and cumulative known cases during an epidemic. When time series data is available for cumulative known cases and some epidemiological characteristics of the infectious disease under study are known, we can estimate the basic reproduction number, the overall reporting rate, the effectiveness of mitigation strategies and the final size of the epidemic. Moreover, from the ODE system, we can determine the epidemic peak and peak time. We apply our model to Canadian national, provincial and territorial cumulative incidence data for the current Covid-19 pandemic to illustrate its predictive power.

2. Model

The epidemic model is a modified SIR model. The compartments we are interested in are the number of mild cases, $I_m$, the number of severe cases, $I_s$, the cumulative known cases $C_K$ and cumulative incidence $C_I$. We make the following assumptions in the model:

- The total population is constant.
- Acquired immunity lasts longer than the epidemic.
- There is no co-infection or super-infection.
- The testing/reporting rate in the population is relatively constant.
- The probability of a case being severe vs. mild is constant.
- All severe infections are reported, whereas a fraction of mild infections are.

The model equations are then given as:

\[
\begin{align*}
\frac{dC_I}{dt} &= R_0 M(t) \left(1 - \frac{C_I}{N}\right) (l_m + p l_s) \\
\frac{dC_K}{dt} &= r (1 - p_s) R_0 M(t) \left(1 - \frac{C_I}{N}\right) (l_m + p l_s) + p_s R_0 M(t) \left(1 - \frac{C_I}{N}\right) (l_m + p l_s) \\
\frac{dl_m}{dt} &= (1 - p_s) R_0 M(t) \left(1 - \frac{C_I}{N}\right) (l_m + p l_s) - l_m \\
\frac{dl_s}{dt} &= p_s R_0 M(t) \left(1 - \frac{C_I}{N}\right) (l_m + p l_s) - l_s
\end{align*}
\]

where $N$ is the total susceptible population, $R_0$ is the basic reproduction number, $p_s$ is the probability of severe infection, $r$ is the average testing/reporting rate, $M(t)$ is a mitigation function, and $t$ is time scaled by the average infectious lifetime $1/\mu$.

Model (1) can be rewritten as:

\[
\begin{align*}
\frac{dC_I}{dt} &= R_0 M(t) \left(1 - \frac{C_I}{N}\right) (l_m + p l_s) \\
\frac{dC_K}{dt} &= r (1 - p_s) \frac{dC_I}{dt} + p_s \frac{dC_I}{dt} \\
\frac{dl_m}{dt} &= (1 - p_s) \frac{dC_I}{dt} - l_m \\
\frac{dl_s}{dt} &= p_s \frac{dC_I}{dt} - l_s
\end{align*}
\]

Note that Model (1) accounts for, but does not assume, that severe infections are less infectious than mild or asymptomatic infections by a factor of $p$. We include $p$ since in novel or dangerous epidemics severe cases are often managed and quarantined once known.

Model (1) is coupled with four initial conditions, derived from three parameters:
where $R$ is the number of cases that have already recovered or died by the time we start fitting data.

Model (1) depends on a pre-defined mitigation function, $M(t)$, that must satisfy certain conditions. We first assume that, in the case of no modified behaviour, $M(t) = 1$. In cases of mitigation strategies taken by populations as a whole, we assume that these strategies have a maximum possible effect, $k$, and that they take time to be implemented to their full effect. In other words, these methods are neither perfect nor immediate. For this study, we prescribe $M(t)$ as

$$M(t) = k + (1 - k)e^{-dt}.$$  (3)

### 3. Data fitting

The goal is to fit Equations (1)–(3) to the number of known infections that are reported by the government or media. That is, we will determine the parameters $R_0$, $r$, $N$, $p$, $p_s$, $\mu$, $d$, $k$, $I_{m0}$ and $I_{s0}$ from a vector of data $\mathbf{C}_k$ i.e., we need to fit the vector of parameters

$$\mathbf{v} = [I_{m0}, I_{s0}, R_0, N, p, r, p_s, \mu, k, d]$$  (4)

to values $\mathbf{v}^*$ such that

$$\mathbf{C}_k(t_{\text{data}}; \mathbf{v}^*) = \mathbf{C}_k$$  (5)

where $t_{\text{data}}$ are the time points associated with $\mathbf{C}_k$ and $C_{m0}$ and $R$ are missing from $\mathbf{v}$. We assume these values are fixed by the data and cannot be fit. Effectively, we assume that there is no error in the number of reported cases and recoveries for day 0 of the fitting routine.

We proceed with a fitting routine that employs the mean-square error, defined as

$$G(\mathbf{v}) = \sum_{t \in t_{\text{data}}} \left( C_k(t; \mathbf{v}) - \bar{C}_k(t) \right)^2 + \left( ( C'_k(t; \mathbf{v}) - \left( \bar{C}'_k(t) \right) ) \right)^2.$$  (6)

Additionally, to fit the model to the data, upper and lower bounds for the vector of parameters must be defined. These bounds are determined by the literature and expert knowledge (Table 1). The least-squares algorithm is quickly implemented in Python through the LMFIT package and it is quick to evaluate, meaning that we can obtain updated estimates in almost real time. In order to fit this model, we take advantage of the fact that solving a small system of ODEs, with a relatively small number of parameters is computationally cheap. We integrate system (1) with the prescribed $M(t)$ and an inital guess at conditions (2). We then use the integrated $C_k(t; \mathbf{v})$ to calculate and minimize the residual.

### 4. Application to Canada’s SARS-CoV-2 outbreak

We illustrate the effectiveness of this model by applying it to the current SARS-CoV-2 outbreak in Canada. We take data of cumulative incidence from the Government of Canada website of Canada (2020). This data set gives the total incidence of

Table 1

| Parameter | Init. Value | Min | Max | Ref. (if available) |
|-----------|-------------|-----|-----|---------------------|
| $p_s$     | 0.1         | Fixed |      |                     |
| $M$       | 0.1         | Fixed |      |                     |
| $R_0$     | 2.5         | 1   | 3.5 |                     |
| $N$       | Fixed       |      |     |                     |
| $R$       | $10^{-8}$   | 0   | 0.8 |                     |
| $P$       | $10^{-4}$   | 0   | 1   |                     |
| $K$       | 0.001       | 0   | 1   |                     |
| $D$       | 0.1         | 0   | 1   |                     |
| $C_{m0}$  | Fixed       |      |     |                     |
| $I_{m0}$  | $10(C_{m0} + 1)$ | $C_{m0}$ | $10(C_{m0} + 1)$ |                     |
| $I_{s0}$  | $0.5(C_{m0} + 1)$ | $C_{m0}$ | $10(C_{m0} + 1)$ |                     |
| $R$       | 1000        | $0.1C_{m0}$ | $N$ |                     |
cases from first reported case at the national and provincial/territorial level. Here-within, we present results for the national level, as well as, two large and two small provinces, Ontario (ON), British Columbia (BC), New Brunswick (NB), and Prince Edward Island (PE), respectively.

2.1. Parameter values and initial conditions

Equations (1)–(3) include eight parameters \( (R_0, r, N, p, p_s, \mu, d, k) \), and four initial values \( (C_{K0}, I_{no}, I_0, \text{and } R) \). We fix two model parameters and one initial condition to reduce the number of function calls to find a fit. We first fix the infectious lifetime, \( \mu \) 10 days (Aylward et al., 2020), as we need the parameter to set our timescale. Second, we assume that the initial number of known cases, \( C_K(0) \) is known and is equal to the first data point that will be considered in the fitting routine. Finally, we fix parameter \( p_s = 0.1 \). Since our total incidence information does not track mild or severe cases, we base this number on the age demography of Canada. Approximately 17% of the population of Canada is 65 years or older (Canada (2020a)). This demographic will make up the majority of severe cases, but will not entirely be affected. Therefore, the value \( p_s = 0.1 \) is reasonable. Additionally, this estimate is consistent with the probability of severe cases reported by the WHO (Aylward et al., 2020). Parameter values and known initial conditions values are listed in Table 1.

For unknown parameter values and initial conditions that will be informed by the model fit, we also define reasonable upper and lower bounds. These bounds are listed in Table 1.

2.2. Cleaning the data

At the beginning of an epidemic, especially in the case of a novel pathogen there are many issues with early reporting data. Events like baseline disease spread, ramping up of testing, panic-induced crowding, etc. can cause the actual disease incidence to be obscured. Even though for some provinces and territories, and at the national level, we have data dating back to January 2020, we cannot use all of it. To determine when testing rates become relatively consistent and the system has ‘settled’ to an extent, we fit eight-day rolling averages of cumulative incidence to a very simple differential equation

\[
\frac{dI}{dt} = R_{eff}I
\]

where \( R_{eff} \) is the effective reproduction number at any given time. These results are shown in Fig. 1 the national scale, and the provinces of Ontario British Columbia, New Brunswick and Prince Edward Island. Note the high variability in the effective reproduction number for the early time points. We therefore eliminate these points and fit the cumulative incidence starting from March 15, 2020. Ignoring these data points can further be justified given that there is a more consistent trend in \( R_{eff} \) estimates after this date. Also, after this date, social distancing measures were put in place, schools were closed, and reporting became daily in all provinces/territories.

2.3. Data fits and projections

Our data set is limited, therefore the typical block bootstrapping used for time series data (7) is not feasible. The method described in (7) breaks the data set into blocks of consecutive time points, and bootstraps over said blocks. Given that we only have approximately 40 data points, we would either have to sacrifice the number of blocks or the number of points in a block, leading to limited success with bootstrapping. To proceed in a manner that provides a mean fit and confidence intervals for sensitivity analysis, we choose to fit the model using random fraction \( P \) of \( M - m \) data points, where \( M \) represents the total number of data points, and \( m = 5 \) is the last \( m \) data points in the set. We reserve the final \( m \) points so that we can observe whether the projected values of the fitted model lie within these data. Here within, we set \( P = 2/3 \). To perform the fits, we also provide initial values for all model parameters. We choose these values randomly for each model fit iteration within their defined bounds (Table 1). We perform the model fitting routine 1000 times for each geographical region under study. Fitted model parameters are shown in Table 2, for Canada, ON, BC, NB, and PE. Note that we fit the national numbers as a whole instead of a sum of parts.

Uncertainty in some of the fitted model parameter values is large. This is a result of a high level of sensitivity of these parameters to the initial conditions, and the \( P(M - m) \) data point distribution used to fit the model. Particularly, we find that limiting the number of data points, especially to fewer than 30 creates a dataset that is too sparse to fit accurately.

Fig. 2 shows the national fit with 95% confidence intervals on a linear and log scale (panels (a) and (b)). Here we see that \( I_m > I_e \) over all time (red vs blue lines). We also see that the total number of cases in Canada is projected to be \( 395005 \pm 1016403 \), where \( 162682 \pm 465543 \) are known/reported. The shaded areas indicate the 95% confidence interval around the mean fitted values. Note that rebound indicated by the red shaded region is an artifact of the sensitivity analysis. When the \( P(M - m) \) data points used to fit the model are clustered to the early data points (mid/late-March and early-April) infection outcome is delayed, giving a later and large infection peak. Instead of ignoring these model fits, we have instead chosen to include the results within the sensitivity analysis to show the sensitivity of the fitting method to the distribution of the \( P(M - m) \) chosen data points.
In panel (c), it is evident that the mean fit (blue line) lies within the red data points, and therefore, the short-term forecasting of the model is useful. We also see that, in panel (d), the reproduction number for Canada is projected to be less than unity in the middle of May 2020. This agrees with other projections in the literature (Daniel et al., 2020). Finally, we note that the rebound indicated in the confidence interval shown in panel (c) is an artifact of the sensitivity analysis, as described previously.

Fig. 3 shows provincial fits with 95% confidence intervals on a linear and log scale for Ontario, British Columbia, New Brunswick and Prince Edward Island. Figs. 4 and 5 show the new known cases per day and $R_{\text{eff}}$ for the aforementioned provinces. Note the long flat peak for the province of Ontario in Fig. 3. This agrees with government of Ontario projections (Tuïte et al., 2020). Again, note the rebounds in cases shown in Figs. 3 and 4. Additionally, we see that these artifacts of the sensitivity analysis increases as the province population size decreases and the number of new cases in the province are small.

From the fitted models, we can determine important epidemiological outcomes at each scale under study. The estimated peak time, peak magnitude, and attack rate for the national scale in Table 3. We also provide the same information for ON, BC, NB, and PE. Due to the nature of the model, the attack rate in this case is the attack rate in the case that current intervention...
measures are maintained until the end of the epidemic. We see the attack rate is relatively consistent across provinces, as many provinces have implemented similar mitigation strategies. Additionally, we see that if the $I_{in}$ peak time differs from the

|               | Canada | Ontario | British Columbia | New Brunswick | Prince Edward Island |
|---------------|--------|---------|------------------|--------------|----------------------|
| $R_0$         | 3.43 ± 0.39 | 3.35 ± 0.64 | 3.30 ± 0.88 | 3.16 ± 1.36 | 2.63 ± 2.20 |
| $p$           | 0.99 ± 0.03 | 0.99 ± 0.12 | 0.98 ± 0.17 | 0.95 ± 0.39 | 0.92 ± 0.43 |
| $r$           | 0.36 ± 0.17 | 0.28 ± 0.25 | 0.38 ± 0.17 | 0.20 ± 0.26 | 0.11 ± 0.26 |
| $N_0$         | 37.6 × 10^6 | 14.4 × 10^6 | 4.97 × 10^6 | 0.78 × 10^6 | 0.15 × 10^6 |
| $k$           | 0.005 ± 0.05 | 0.02 ± 0.21 | 0.03 ± 0.19 | 0.05 ± 0.32 | 0.19 ± 0.7 |
| $d$           | 0.11 ± 0.02 | 0.11 ± 0.14 | 0.23 ± 0.13 | 0.51 ± 0.30 | 0.32 ± 0.43 |
| $I_{in}$      | 3461 ± 1460 | 973 ± 804 | 999 ± 427 | 46 ± 58 | 29 ± 60 |
| $I_{0}$       | 221 ± 112 | 86 ± 52 | 43 ± 26 | 2 ± 5 | 2 ± 5 |
| $C_{00}$      | 253 | 103 | 73 | 2 | 1 |
| $R_0$         | 279 ± 2861 | 7280 ± 194802 | 1790 ± 55768 | 2679 ± 29221 | 1256 ± 8692 |

Fig. 2. Model fits at the national level. The mean values and 95% confidence interval of all model populations are shown on a linear and log scale (panels (a)–(b)). Panels (c) and (d) show the resulting number of new cases per day, and the national value of $R_{eff}$ over time. (top rows) mild cases $I_{m}$ (red), severe cases $I_{s}$ (blue), known cases $C_{k}$ (green), total cases $C_{i}$ (green). Panels (a)–(c) dark blue and red points are the known data. In all plots the shaded area shows 95% confidence intervals around the mean.
Fig. 3. Mean fits of model (1) to cumulative case data for ON, BC, NB, and PE from left to right, shown on a (top row) linear and (bottom row) log scale. Mild $I_m$ (red) and severe $I_s$ (blue) cases are shown, as well as, known cases $C_k$ (green) and the total number of infections $C_I$ (purple). Shaded area shows 95% confidence intervals. The dark blue and red points indicate the case data.
Fig. 4. The number of new cases per day for ON, BC, NB, and PE (left to right). Subplots (e), (f), (g) and (h) are the same plots without the confidence intervals for clarity. The mean of the model fits is shown (blue line), with 95% confidence interval (shaded area) in the top row. The dark blue and red point indicate the case data.
If peak time, then $I_m$ peak time precedes the $I_s$ peak time. We also report the peak time of $C_K$ which represents the number of known new cases per day. The peak number of known new cases always precedes the actual peak. This should be expected, and is likely exaggerated because the model assumes cases are reported and known as soon as they present.

**Table 3**
Table of epidemiological outcomes for Canada, ON, BC, NB, PE.* (The first peak of each fit is reported as consecutive peaks are artifacts of the sensitivity analysis.)

|                  | Canada            | Ontario           | British Columbia | New Brunswick | Prince Edward Island |
|------------------|-------------------|-------------------|------------------|---------------|---------------------|
| **Peak Magnitude** |                   |                   |                  |               |                     |
| $I_m$            | $44574 \pm 20874$ | $31881 \pm 74334$| $1149 \pm 15286$| $152 \pm 201$ | $50 \pm 61$         |
| $I_s$            | $4952 \pm 2319$   | $3542 \pm 8259$  | $112 \pm 1696$  | $16 \pm 22$   | $5 \pm 5$           |
| $I = I_m + I_s$  | $49527 \pm 23194$ | $35423 \pm 82593$| $1258 \pm 16965$| $168 \pm 224$ | $55 \pm 65$         |
| **Peak Time**    |                   |                   |                  |               |                     |
| $I_m$            | 2020–05 – 14      | 2020–06 – 03      | 2020–03 – 27     | 2020–04 – 03  | 2020–04 – 01        |
| $I_s$            | 2020–05 – 14      | 2020–06 – 03      | 2020–04 – 01     | 2020–04 – 03  | 2020–04 – 02        |
| $I = I_m + I_s$  | 2020–05 – 14      | 2020–06 – 03      | 2020–03 – 28     | 2020–04 – 03  | 2020–04 – 01        |
| $C_K$            | 2020–05 – 04      | 2020–05 – 25      | 2020–03 – 19     | 2020–03 – 26  | 2020–03 – 24        |
| **Attack rate (C/| 0.01 ± 0.02       | 0.03 ± 0.17       | 0.003 ± 0.07     | 0.01 ± 0.13   | 0.01 ± 0.09         |

Fig. 5. Model results for the effective reproduction number for ON, BC, NB and PE. Shaded area shows the 95% confidence interval around the mean (blue line).
3. Discussion

We propose a simple epidemiological compartmental model that can be readily fit to cumulative incidence data. The goal of the model is to synthesize disease dynamics to the fewest possible parameters while remaining robust enough to be able to project possible outcomes and to provide estimates that can help inform policy and healthcare needs assessments. The intent of this model is to be able to get a rough idea of epidemic trajectory in cases with limited data.

The model itself is unique in that it estimates a number of key epidemic parameters without the assumption that the growth of the known cases is directly proportional to the total number of cases. By separating out known cases from actual cases, we can better estimate parameters such as the true \( R_0 \) and attack rate. Our estimates for the peak time of the epidemic is decoupled from the peak testing/reporting rate, giving a much better idea of when the epidemic might peak. Of note, qualitatively the model suggests that the peak in known cases will precede the actual epidemic peak.

Our estimates for \( R_0 \) are consistent with those already estimated in the literature (Aylward et al., 2020; Lai, Bergna, et al., 2020; Lai et al. 2020; Liu et al., 2020; Tang et al., 2020), and the peak times estimated are consistent with the timelines of other countries for which the infection has peaked, such as Italy (Fanelli & Piazza, 2020).

We are also able to effectively estimate a number of parameters which, when compared among regions can help inform why a similar epidemic may be spreading differently in two different populations, or which mitigation strategies are most effective. The parameter \( p \), for instance, can be interpreted as the efficacy of isolation of severe cases and how quickly isolation is implemented for known severe infections. The parameter \( r \) can be used to compare testing and reporting rates among different regions. For instance, in our fits we see that British Columbia has a higher reporting/testing rate (a greater value for \( r \)) as well as a quicker identification of severe cases (lower \( p \)) than Ontario which may point to how BC has been able to control the spread of SARS-CoV-2 more effectively than ON.

The model has its drawbacks; mostly related to the fitting. Fitting ODEs to parameters is generally sensitive to the initial choice of parameters. We have done our best here to prescribe parameters that will work for SARS-CoV-2 cumulative incidence curves, but care should be taken when implementing that the fits are indeed viable. While least-squares is very quick to implement and analyze, it does not allow for strict confidence intervals. More robust methods, such as basin-hopping may allow for better parameter estimates. Additionally a hierarchical fit, when all provinces and territories are fit with individual but related parameters, as well as, a national fit may allow for better parameter estimates. However, here, distribution characteristics that relate a specific model parameter for each fit would then also have to be assumed i.e., assume that in all provinces and territories the fraction of severe cases will be related on a normal distribution with a mean and standard deviation. However, this is difficult to do since (1) many of the parameters are affected by different provincial and territorial government policies i.e., reporting rate, mitigation strategy, and (2) the distribution parameters (i.e., mean and standard deviation) would also have to be estimated alongside the model parameters.

Improvements to the model can also be added. The reporting rate \( r \), for example, could be assumed to change over time, making \( r = r(t) \). Additionally, different mitigation function structures, could be tested. Finally, some provincial and territorial COVID-19 outbreaks are dominated by imported cases (from travelling) rather than community transmission. The model could be modified to include this distinction to determine contact tracing rates by geographical region. However, this again, would depend on data quality and availability.

Declaration of competing interest

The authors have no conflicts of interest arising from or with respect to this study.

Acknowledgments

This research is supported by NSERC Discovery Grant 2019-05679.

References

Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. PloS One, 15(3), Article e0230405.

Aylward, R., Liang, W., Dong, X., Eckmanns, T., & Fisher, D. (2020). Report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). Canada, S. (2020a). Table 17-10-0005-01 Population estimates on July 1st, by age and sex. Canada, S. (2020b). Table 17-10-0009-01 Population estimates, quarterly. Statistics Canada.

Daniel, C., Frid, L., Rayfield, B., & Embrey, A. (2020). A scenario-based modeling framework for projecting COVID-19 infections and deaths: Case study application for Canada. Fanelli, D., & Piazza, F. (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos. Solitons & Fractals, 134, 109761.

Fang, Y., Nie, Y., & Penny, M. (2020). Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. Journal of Medical Virology, 92(6), 645–659.

Fisman, D. N., Hauck, T. S., Tuite, A. R., & Greer, A. L. (2013). An IDEA for short term outbreak projection: Nearcasting using the basic reproduction number. PloS One, 8(12).

Lai, A., Bergna, A., Acciarri, C., Galli, M., & Zehender, G. (2020a). Early phylogenetic estimate of the effective reproduction number of SARS-CoV-2. Journal of Medical Virology, 92(6), 675–679.

Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J., & Hsueh, P.-R. (2020b). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): The epidemic and the challenges. International Journal of Antimicrobial Agents, 105924.
Liu, Y., Gayle, A. A., Wilder-Smith, A., & Rocklov, J. (2020). The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine, 27*(2) (published online).

Government of Canada Covid-19 data. (2020). *Coronavirus disease (COVID-19): Outbreak update*. https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection.html#a1.

Tang, B., Xia, F., Tang, S., Bragazzi, N. L., Li, Q., Sun, X., Liang, J., Xiao, Y., & Wu, J. (June 2020). The effectiveness of quarantine and isolation determine the trend of the COVID-19 epidemics in the final phase of the current outbreak in China. *International Journal of Infectious Diseases, 95*, 288–293.

Tuite, A. R., & Fisman, D. N. (2018). The idea model: A single equation approach to the Ebola forecasting challenge. *Epidemics, 22*, 71–77.

Tuite, A. R., Fisman, D. N., & Greer, A. L. (2018 Mar 1). Mathematical modelling of COVID-19 transmission and mitigation strategies in the population of Ontario, Canada. *Canadian Medical Association Journal, 22*, 71–77.

Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., Cuomo-Dannenburg, G., Thompson, H., Walker, P. G., Fu, H., & others. (June 2020). Estimates of the severity of coronavirus disease 2019: A model-based analysis. *The Lancet Infectious Diseases, 20*(6), 669–677.