A scenario-based modeling framework for projecting COVID-19 infections and deaths: Case study application for Canada

DRAFT

Colin Daniel, Leonardo Frid, Bronwyn Rayfield & Alex Embrey
Apex Resource Management Solutions
www.apexrms.com

April 20, 2020

Introduction

Decisions are being made every day, at the level of individuals to countries, in response to the outbreak of COVID-19. Ideally these decisions are informed by the best possible forecasts – forecasts that provide answers to questions such as: how many new infections and deaths can I expect in my city/province/country tomorrow? In two weeks time? By the end of the summer? The better our forecasts, the better we can prepare for the future – whether it be mobilizing health care resources to deal with the anticipated number of future infections, or understanding, in advance, the potential consequences of changes in public health measures. While the need to forecast disease progression is nothing new, the speed and extent of the current COVID-19 pandemic has challenged our abilities, as forecasters, like never before. Early data on the disease's epidemiology is still very limited; records of cases and infections are far from complete. And in most parts of the world the dynamics of the disease are still changing daily. Yet despite these challenges, decision makers still need to push forward and make new decisions every day.

There are several different approaches to forecasting present and future levels of disease infections and deaths. Models can be as simple as estimating the future number of cases and deaths based on published infection and fatality rates. More complicated mathematical models, however, can also be used. Two of the most widely cited COVID-19 mathematical models to-date include the U.K.'s Imperial College London model (Flaxman et al. 2020), and the U.S.-based IMHE COVID-19 model (Murray 2020).

A common feature of these two models is that they both rely on reported deaths, rather than cases, in order to generate their projections. This is due to the well-known challenges associated with the collection of case data. For most jurisdictions, deaths are more much more accurately reported than cases, particularly in the early stages of an outbreak where testing is far from random. As a result, a model developed based on death data is generally much more reliable than one generated using cases (the exception, of course, being when case data is randomly sampled). The challenge with using death data for projecting infections, however, is that deaths generally lag cases by 1-2 weeks, making death-based projections somewhat less responsive to rapidly changing conditions. A second common feature of both these models is their stochastic nature. Both models can incorporate statistical uncertainty in model input parameters into their projections, thus producing ranges around the resulting projections.

Beyond their shared reliance on death data, however, the Imperial College and IMHE models differ in their mathematical approach. Simply put, the Imperial College model uses a traditional SIR (Susceptible-Infected-Recovered) approach, while the IMHE model instead fits a statistical growth curve model to death data. Each approach has its strengths and weaknesses. The SIR approach generally produces better medium to long-term projections, as it can account for the change in a population from Infected

to Recovered as the disease progresses. However an SIR model requires more model parameters than a growth curve model, some of which can be challenging to estimate in the early stages of an epidemic. By contrast a model fitted to death data can perform well in the early stages of an epidemic, where data is sparse and few people have yet to be infected; it's limitation is that it can only forecast forward for as long as relevant reference death data exists in other reference jurisdictions from which to estimate each jurisdiction's future growth curves.

An important limitation of both models, however, is their prescriptive nature: forecasts are "published" by a team of modelers, with the modeling scenarios decided upon *a priori* by the modeling team. For example the results of the IMHE model are published regularly for all U.S. states (and some countries) under single set of scenario assumptions; as a result there is no opportunity for policy makers in these jurisdictions to decide for themselves what scenarios to consider. Similarly, the first set of COVID-19 results from the Imperial College model considered a specific set of historical public health interventions in 11 European countries. This approach can work well for policy makers equipped with their own personal team of modelers. However for many jurisdictions it is simply not possible to support such a modeling team.

This paper presents an alternative approach to projecting COVID-19 infections and deaths. Here we have developed a general *framework* for modelling infections and deaths, rather than a specific model, that can be configured for use with any jurisdiction, of any size, in the world. By framework we mean that we allow end users to develop their own model, specific to their jurisdiction and questions. This modelling framework, called SyncroSim, is designed specifically for the rapid deployment of complex open-source simulation models. SyncroSim has evolved from our experience over the past 25 years developing models for forecasting animal populations and vegetation change, two areas in which, like the early stages of the COVID-19 outbreak, there is typically very high uncertainty and very little data. While we demonstrate the framework here with a specific COVID-19 model of infections, the framework we present here can easily be adapted to other forms of disease simulation models.

The model for COVID-19 that we use here to demonstrate our modeling framework is similar to the approach used by the IMHE model, in that we fit growth rate curves to various time series of COVID-19 death data. Unlike the IMHE model, however, we also track infections in our model. In addition to actual death data, our model requires users to make assumptions regarding future growth rates of COVID-19 deaths, including changes in growth rates under alternative public health interventions. As we will demonstrate, such growth rate "scenarios" are typically developed through an analysis of existing death data and/or data and projections from other jurisdictions. Other model inputs include the infection fatality rates, attack rate, and infection period, for which initial estimates are available in the published literature.

The modelling framework is stochastic, in that it allows uncertainties to be specified for model inputs; this, in turn, allows the consequences of these uncertainties to be reflected in projections of infections and deaths. Importantly, because we are using a framework, rather than a specific model per se, most of the assumptions within the model can be modified by users as they see fit. Model equations can also be reviewed and modified, as required, and the code associated with the actual model equations is open-source. Finally, and we believe most importantly, the framework is scenario-based, allowing users to easily game with alternative "what-if" scenarios regarding model structure, model assumptions, and the timing and extent of future public health measures.

In the remainder of this paper we describe the approach used by this framework and the software that underlies its development; we finish with a demonstration of how the framework can be applied to develop a model for use in Canada.

Model Description

Calculating infections and deaths

Given an incubation period (C_t) and symptom-to-death period (S_t) for each day t in our simulation, the model can determine the total infection period each day as $i = C_t + S_t$. The model then uses a time series of actual daily death data, $\{D_t : i + 1 \leq t \leq i + n\}$, for days $i + 1$ to $i + n$, to back-calculate the prior cumulative number of infections, $\{I_t : 1 \leq t \leq \text{n}\}$, for days 1 to n, based on the relationship between daily infections and daily deaths, namely $\Delta I_t = I_t - I_{t-1} = D_{t+i} / F_t$, where F_t is the infection fatality rate. This is repeated for all days t, $1 \leq t \leq n$, for which death data exists. For example, assuming an infection fatality rate 1%, along with an incubation period 5 days and a symptom-to-death period 15 days, the model would calculate that a single death occurring 20 days from now would correspond to 1 / 0.01 = 100 daily infections today. Note that the variables C_t , S_t and F_t are all represented as discrete-time stochastic processes – in other words, they are all random variables that can vary over time; as a result the projected cumulative daily infections, I_t , is also a stochastic process.

The model then uses a simple time series growth model to project the daily cumulative infections forward beyond the last day, n, for which infections can be directly estimated from deaths (i.e. to project I_t for $t > n$). Currently two forms of growth models are supported: exponential and logistic. If the exponential growth model is selected, the future number of daily cumulative infections is calculated as $I_{t+1} = I_t (1 + R_t)$, where R_t is growth rate for day t. A limit to the total number of cumulative infections can also be set as A_tP, where A_t is the user-defined attack rate on day t and P is the total population size. By contrast if the logistic growth model is selected, the future daily cumulative infections is calculated using the discrete form of the logistic equation: $I_{t+1} = I_t R_t ((1 – (I_t – 1 / A_tP)) + 1)$, where R_t is the user-supplied maximum growth rate for day t, and A_tP represents the logistic model's traditional carrying capacity. Note that these two model forms were selected as they both can be parameterized using at most two parameters: a growth rate and an optional attack rate; additional model forms could be easily added in the future (Figure 1). Note also that, as with the other model variables outlined above, the growth rate (R_t) and attack rate (A_t) are also both discrete-time stochastic processes, and thus can also be specified as distributions that vary over time. Finally, the future number of daily deaths, D_t, for those days $t > n$ for which deaths are not provided as a model input, are calculated from the projected number of daily cumulative infections as $D_t = \Delta I_{t-i} F_{t-i}$.

Figure 1 contrasts the two different forms of growth models that are currently available in the model. With the logistic model the growth rate is exponential at low population levels, but then declines over time as the cumulative number of infections approaches the product of the population and attack rate. By contrast, with the exponential model the cumulative number of infected increases exponentially at a fixed daily growth rate until the maximum possible infections are reached. However because the growth rate can optionally vary over time, it is possible to create a third type of model, which we refer to as "time-varying" (or non-stationary) exponential growth model, whereby the growth rate of the exponential model changes over time – e.g due to a gradual reduction in growth rate as a result of the introduction over time of public health measures. It is this final option that provides the greatest

flexibility for specifying future patterns of infection growth rates in the model, as ultimately any empirical pattern of infection growth can be captured in this way.

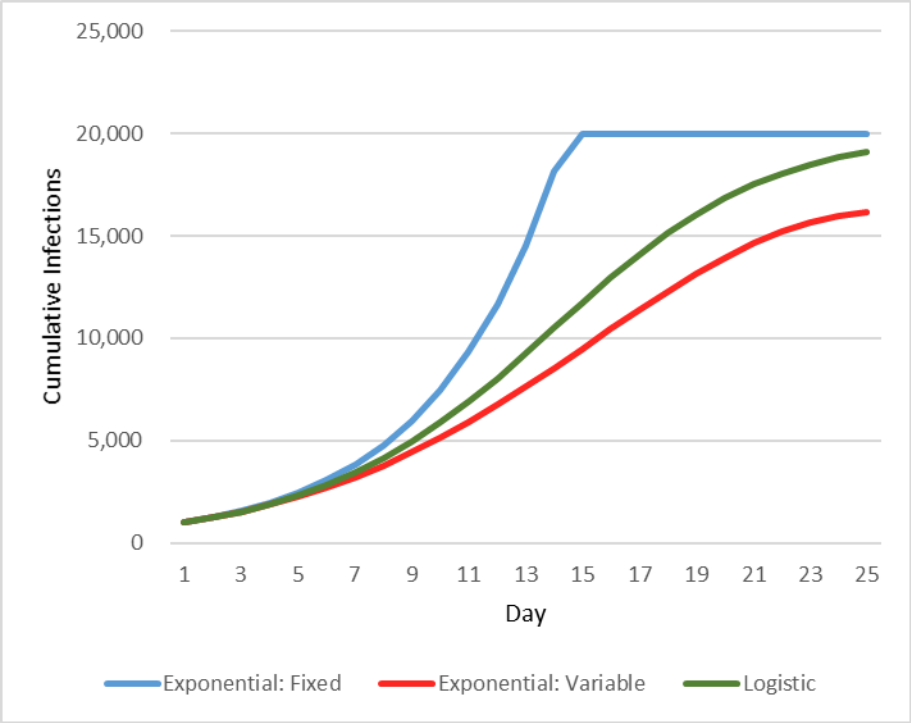

Figure 1: Cumulative infections projected by the exponential and logistic forms of the infection growth model. Example shows deterministic projections over time starting with 1000 infections, growth rate of 25%, population of 40,000, and attack rate of 50%. The variable exponential curve lowers the initial growth rate of 25% by 1% every day, demonstrating how time-varying growth rates can be used to capture the effects of public health measures on cumulative infections.

Parameterizing and running the model

The equations in the previous section detail the calculations required for a single time series projection of daily infected and deaths for one jurisdiction. The model framework is designed to allow users to run multiple jurisdictions simultaneously (e.g. several counties within a state, or several countries within a region), whereby some inputs are shared by all jurisdictions (e.g. incubation and symptom-to-death periods), while others can vary between jurisdictions (e.g. deaths, population and fatality rates). The framework is also designed to allow some variables (e.g. growth rates) to vary over time; the choice of which variables to vary by jurisdiction and/or date is fully customizable by the user. Finally, a key element of the SyncroSim modeling framework is its ability to capture uncertainty in projections through the use of random variables and Monte Carlo simulations. As in all Monte Carlo simulations the model can be run for multiple Monte Carlo realizations, with the results of each realization differing from each other due to the random sampling of input values for those model inputs that are uncertain.

As shown in Table 1, most of the model parameters can be specified as distributions, rather than single values; these distributions can also vary spatially (i.e. by jurisdiction) and temporally (i.e. by day). In this way the model can capture a wide range of uncertainty and variability in model parameters.

With the support of the SyncroSim's modeling framework, users are able to chose from a wide range of distributions in order to characterize the uncertainty in model inputs, including empirically generated frequency distributions. Users are also able to vary these distributions over time, and by jurisdiction (e.g. when running a multi-jurisdictional model that might divide a country into several states or provinces). This is one of the key benefits of developing a simulation model using SyncroSim: the framework manages all the Monte Carlo simulations, feeding the required sampled inputs to the core model for each realization and timestep.

The final output of the model is a set of model realizations for the projected time series of daily infections and deaths, which can either be post-processed externally (e.g. using scripts written in R with the help of the rsyncrosim R package), or summarized directly within the SyncroSim platform into time series of mean and Monte Carlo confidence intervals for daily and cumulative infections and deaths.

Table 1: List of the model inputs and outputs, including which variables can be specified by jurisdiction, date and/or Monte Carlo realization. Input variables that vary by realization can be characterized using either parametric distributions or user-specified empirical frequency distributions. Parametric distributions currently supported by SyncroSim include uniform, normal, beta and gamma.

Type	Variable Name	Equation Symbol	X = required o = optional		
			Jurisdiction	Date	Realization
Input	Population	P	X		
	Deaths	D_t	X	X	
	Growth rate	R_t	o	o	o
	Attack rate	A_t	o	o	o
	Fatality rate	F_t	o	o	o
	Incubation period	C_t	o	o	o
	Symptom-to-death period	S_t	o	o	o
	Model type		o	o	
Output	Daily Infections	ΔI_t	X	X	X
	Daily Deaths	D_t	X	X	X

The source code for the current model is available in open-source format on GitHub at https://github.com/ApexRMS/epidemic. To streamline development, the first version of the core model is written in C# – doing so allowed us to finish and distribute a fully functional version of the model in under two weeks. As SyncroSim supports models written in any language the model could easily be translated to other programming languages (e.g. R or Python) in the future.

Case Study: COVID-19 in Canada

The best way to understand the model is to see it in action – that is, to see how it might be applied to make real-time COVID-19 projections. For this case study we chose to focus on our home country of Canada, demonstrating the model for five jurisdictions: the country as a whole, along with the four largest Canadian provinces by population, namely Alberta, British Columbia, Ontario and Quebec (representing 99% of the country's COVID-19 deaths to-date). To reiterate the point made earlier

regarding the generality of the framework, the model presented here can easily be configured to run with any set of jurisdictions, ranging from a single public health unit to multiple countries; the framework could also be configured for diseases other than COVID-19.

It is important to note that the model presented here is simply a demonstration of the model framework, and should not be considered actual projections for any of these jurisdictions. We are modelers, not epidemiologists, and so are in no position to make actual COVID-19 projections. Our intent here is simply to demonstrate how the SyncroSim modeling framework might help others in the future with COVID-19 projections for their specific jurisdictions. Note also that the projections presented here are a one-time snapshot, based on the data available at the time of writing this document (i.e. COVID-19 death data publicly available up to and including April 18, 2020). As you will see, however, the framework is designed to support real-time daily updates to its forecasts.

As background, Figure 2 provides an overview of the progression of COVID-19 across Canada and its four major provinces, highlighting strong regional differences across jurisdictions with respect to the rate of spread and number of deaths.

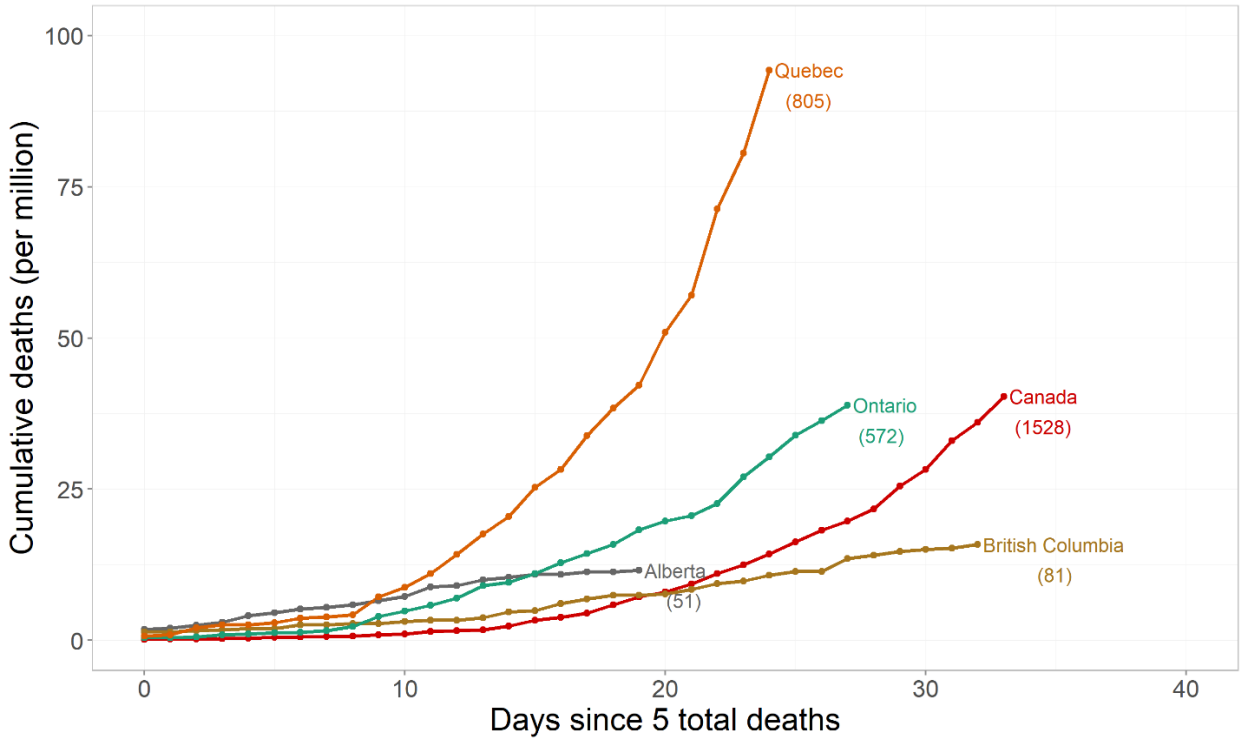

Figure 2: Per capita cumulative deaths due to COVID-19 in Canada, shown for the country as a whole and four provinces, as reported to April 18, 2020. Total reported deaths for each jurisdiction are shown in brackets. Source: Berry et al. (2020).

Our goal in this Canadian case study is to parameterize our COVID-19 model, using the daily death data shown in Figure 2, in order to project, for each jurisdiction: (1) the cumulative number of infected as of today and, (2) the daily number of infected and corresponding deaths into the future. While the model can, in theory, simulate forward in time for any number of days, we have purposefully limited our future projections to a two-week window forward from today; the rationale for selecting this window will

hopefully become evident as we step through the model parameterization in the remainder of this document. Note, however, that a two-week forward projection from today is equivalent to a 5-week forward projection for infections, given that deaths reported today correspond to infections that occurred approximately three weeks prior.

In this case study we developed two possible "what-if" future scenarios: (1) a *Current Measures* scenario, in which we project the consequences of the actual public health measures enacted in each jurisdiction; for the purposes of this demonstration we assume that public health measures began in all our Canadian jurisdictions on March 14, 2020, and (2) a *No Control* scenario, in which we modelled what might have occurred in each jurisdiction had no public health measures been taken on March 14. We also ran each of these two scenarios using two different assumptions regarding the infection fatality rate: a first option in which we used published rates (*Base Fatality* sub-scenario), and a second option in which we increased published rates by 50% (*High Fatality* sub-scenario). This second option was included to explore the sensitivity of our projections to our initial base assumptions regarding the fatality rate, in an effort to reflect the uncertainty regarding the applicability of standard fatality ratios to jurisdictions in Canada at present (e.g. due to the high number of deaths that have occurred to date in long-term care facilities for the over 80 age class).

Case study model inputs

As outlined in the *Model Description* section above, there are several inputs required in order to run the model; a summary of the values assigned to each of these inputs for the case study is provided in Table 2.

Table 2: Model inputs for the Canadian case study.

Model Input	Scenario	Values	Approach	Data Source
Population	All	Reported population by jurisdiction	As reported	Statistics Canada (2020)
Deaths	All	Reported daily deaths by jurisdiction	As reported	Berry et al. (2020)
Growth rate	No Control	Fixed rate by jurisdiction: CA: 21.8% AB: 20.7% BC: 9.6% ON: 27.8% QC: 24.6%	Log-linear regression by jurisdiction of cumulative deaths on or before 2020-04-05	Berry et al. (2020)
	Current Measures	Sampled rates by jurisdiction, date & realization	7-day moving average of death growth rates sampled from Italy, France, Spain and South Korea	Dong et al. (2020)
Attack rate	All	Single distribution: 50-80%	Uniform distribution over published range	Verity et al. (2020)
Fatality rate	Base Fatality	Distribution by jurisdiction: CA: 1.02% (0.55-1.96) AB: 0.80% (0.43-1.55) BC: 1.07% (0.58-2.07) ON: 1.01% (0.55-1.94) QC: 1.11% (0.60-2.13)	Age standardized fatality rate for each jurisdiction, based on published infection fatality ratio (with 95% credible intervals) and age-structured population for each jurisdiction. Fitted to gamma distribution.	Verity et al. (2020) Statistics Canada (2020).
	High Fatality	As above X 1.5	Base rates increased by 50%	Verity et al. (2020)
Incubation period	All	Single distribution: 4.5-5.8 days	Uniform distribution over published 95% credible interval	Lauer et al. (2020)
Symptom-to-death period	All	Single distribution: 16.9-19.2 days	Uniform distribution over published 95% credible interval	Verity et al. (2020)
Model Type	No Control	Logistic (with fixed growth rate)		
	Current Measures	Exponential (with variable growth rate)		

The first class of inputs are those that characterize the behaviour of the disease in general, namely the incubation period (i.e. time from infection to first symptoms), and the duration from onset of symptoms to death. Estimates for the distribution of incubation periods for COVID-19 were taken from Lauer et al. (2020), while estimates for the distribution of onset-of-symptoms-to-death periods were taken from Verity et al. (2020). For this demonstration we assumed these values are the same across all jurisdictions in Canada, and that these values do not vary over time. We captured uncertainty in these values by sampling once for each Monte Carlo realization from the distribution of reported values for each of these two periods (Table 2).

The infection fatality rate was calculated for the *Base Fatality* sub-scenarios using age-specific infection fatality ratios (Verity et al., 2020). These rates were then age-standardized based on the age structure of each jurisdiction's population. Credible intervals (95%) were also estimated for each jurisdiction using age-weighted averaging of the original published age-specific credible intervals. The mean and credible intervals for each jurisdiction were then fit to a gamma distribution and sampled once for each realization. For the *High Fatality* sub-scenarios we increased the mean of the base gamma distributions by 50% (yet left the standard deviation unchanged).

A logistic growth model was used to project the future number of infections under the *No Control* scenario. The *No Control* scenario relies on an estimate of the maximum number of possible infections in the population in order to eventually limit the total number of people infected, and thus requires an estimate of the attack rate. For all modeled jurisdictions we used the estimate of 50-80% for the attack rate (Verity et al., 2020), sampling from this range, using a uniform distribution, once for each realization. The maximum growth rate for the *No Control* logistic model was calculated separately for each jurisdiction as the average growth rate of cumulative deaths over the period before any major public health measures began – i.e. from the day 5 deaths were first reported until 23 days after significant public health measures were enacted (where 23 days represents the average total infection period). We used growth rates of deaths as a proxy for growth rates of infections, which thus assumes fatality rates and infections periods are not changing significantly over the period of our analysis. March 14, 2020 was selected as the last day prior to public health measures for all jurisdictions, corresponding to a cut-off date of April 5, 2020 (i.e. 23 days later) for death data used to estimate the *No Control* growth rate. The average growth rate was then estimated using a log-linear regression on the time series of cumulative deaths over this period for each jurisdiction. All regression models showed excellent fit (R^2 values ranging from 0.93-0.99), with *No Control* estimates of daily growth rates provided in Table 2 above.

For the *Current Measures* scenario we used an exponential growth model with a time-varying (i.e. non-stationary) growth rate in order to project the future infections each day. Daily reported deaths in Canada, by province and date, were taken from Berry et al. (2020). Growth rates for the *Current Measures* scenario were estimated for the period beyond which infections cannot be estimated directly from deaths. With death data available at the time of writing through until April 18, this corresponds to the growth in infections beginning on March 27 (i.e. 23 days prior). While there are many different methods that could be used to estimate a distribution for future growth rates for each of our modeled jurisdictions, for our demonstration we chose to base our future growth rates on trends observed for other "reference" jurisdictions – i.e. jurisdictions that had enacted public health measures similar to those used in Canada, yet were ahead of Canadian jurisdictions with respect to the date upon which these measures were taken. Figure 3 shows the 7-day moving average of cumulative death growth rates for several jurisdictions, both in Canada and internationally, adjusted to show the progression in death growth rates relative to when outbreaks began in each jurisdiction. With the exception of the United States, there is a consistent trend across both the Canadian and international jurisdictions with respect to the decline in growth rates over time. Italy and South Korea are the furthest advanced in terms of time since the start of outbreak, with both countries approximately 3 weeks ahead of Canada in this regard, while France and Spain are about 2 weeks ahead.

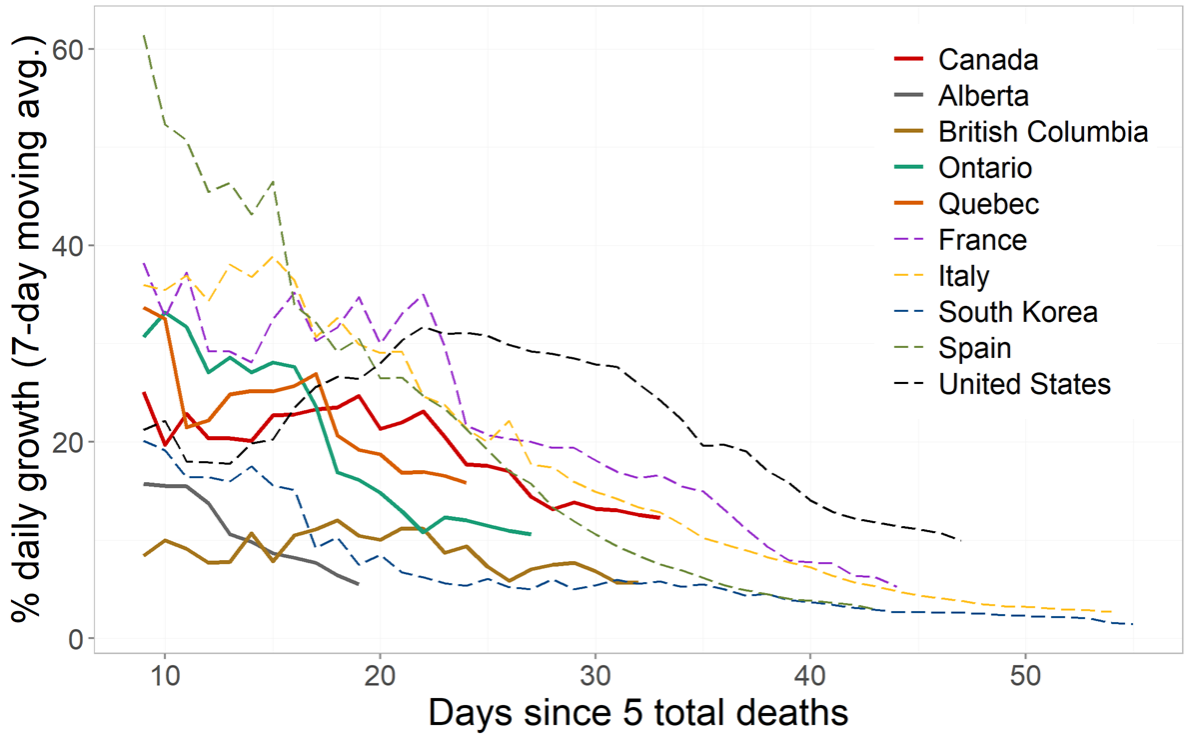

Figure 3: Seven-day moving average of the daily growth rates in cumulative COVID-19 deaths for Canada and four provinces (solid lines), as compared to 5 other countries (dashed lines). Growth rates for each country are displayed relative to the first day in each jurisdiction for which 5 total deaths were reported. Data shown is current to April 18, 2020. Sources: Berry et al. (2020) & Dong et al. (2020).

For our demonstration we selected four countries – Italy, Spain, France and South Korea – to act as reference jurisdictions for our projections. For each of our modeled jurisdictions we found the first point in time on each reference country's 7-day moving average growth curve that was equal to the last recorded 7-day moving average growth rate for our modeled jurisdiction. The reference jurisdiction's growth rate beyond this point in time was then used to extrapolate the future growth rate each day for the modeled jurisdiction, with the last growth rate value for each reference country held constant for all days beyond the end of the reference growth rate time series. Figure 4 illustrates the use of this extrapolation technique for the *Canada* jurisdiction, resulting in four possible future growth rate trajectories for this modeled jurisdiction. We repeated this same approach for all five of our modeled jurisdictions.

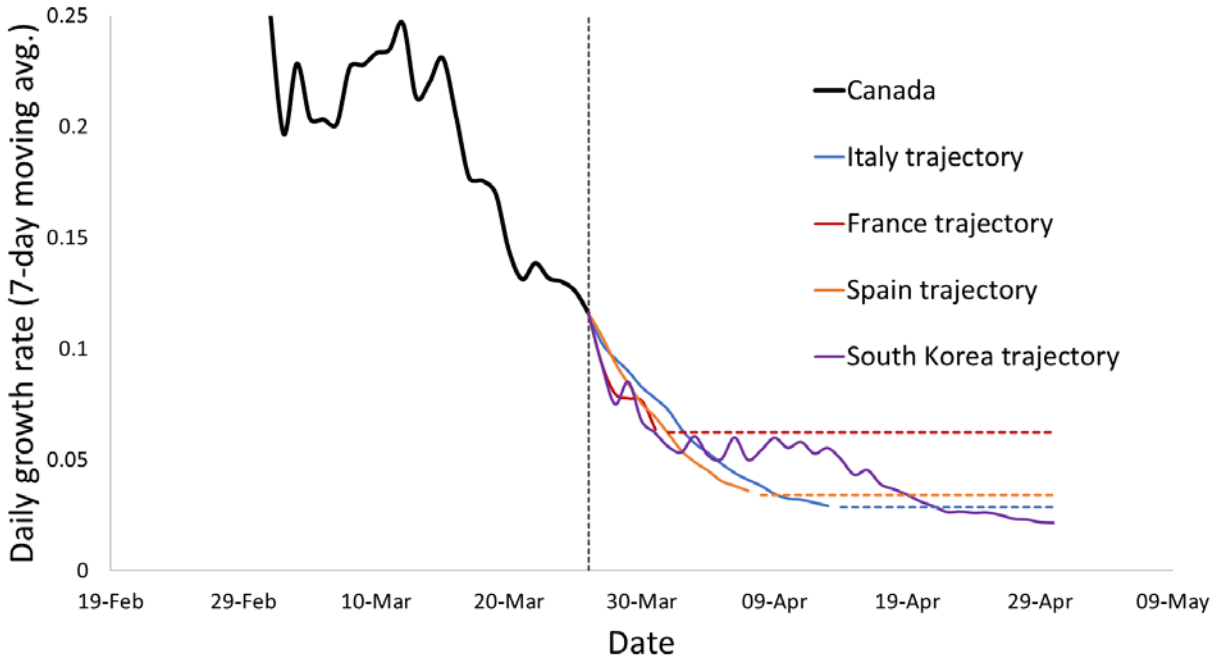

Figure 4: Growth rates in cumulative COVID-19 infections, as supplied as inputs to the model for the *Current Measures* scenario for the modeled *Canada* jurisdiction. Black line represents actual death growth rates calculated from a 7-day moving average of the recorded cumulative deaths, shifted back 23 days from when deaths were recorded. Dotted line indicates the last day for which infections in the model were projected directly from death data; as a result the data represented by the black line are not directly used in the model. Solid coloured lines indicate the trajectory of growth rates corresponding to each reference jurisdiction, based on 7-day moving average of actual growth rates and shifted in time so as to continue the trajectory of the modeled jurisdiction's actual growth rate trend. Dotted lines indicate those reference trajectories that were held constant over time beyond their last actual growth rate value. Data shown are current to April 18, 2020. Sources: Berry et al. (2020) & Dong et al. (2020).

Given the uncertainty that exists as to which of these reference trajectories each of our modeled jurisdictions will follow most closely in the future, we set up our simulations to sample one of these four curves for each of our realizations, each with equal probability. This provides the model with an empirical distribution of future growth curves that spans the full range of possible reference jurisdiction growth rates. Note that this algorithm for estimating future growth rates is particularly powerful because it self-updates the distribution of projected growth rates every day based on the latest death data from each modeled jurisdiction and the four reference jurisdictions; the algorithm also allows reference jurisdictions to be easily added and removed as conditions change in the future.

The start date for all of our case study simulations was set to February 12, 2020, 25 days before the first reported death in Canada; we chose this date in order to capture all of the projected infections that might result through the stochastic back-calculation of infections from deaths. Each of our five modeled jurisdictions were initialized with their actual death data up to and including April 18, 2020. Simulations were run to May 2, 2020 (i.e. for a total of 81 days), two weeks beyond the last day for which deaths were reported. All simulations were repeated for 1000 Monte Carlo realizations.

Case study model results

Figures 5-7 show the results of our demonstration simulations, for each of the five jurisdictions, under the *Current Measures* scenario (assuming the base fatality rate); projections are also summarized in Table 3. Based on the parameterization outlined in the previous section, the model projects an increase in the cumulative number of deaths for Canada from the actual total of 1528 on April 18 to between 3700 and 4250 on May 2 (based on the range of this scenario's 95% Monte Carlo confidence interval). This represents a projected average daily growth rate in deaths of between 6.5-7.6% for the entire country over this two-week period. The demonstration model, in turn, projects a corresponding cumulative number of infections for the entire country ranging between 265,000 and 1.2 million (mean=594,000) on April 18, rising to between 400,000 and 2 million (mean=973,000) by May 2.

The demonstration model also captures regional differences in projections between provinces. Alberta and British Columbia are projected to have average daily growth rates in deaths between April 18 and May 2 of 3.3-5.4% and 3.1-5.1%, respectively, as compared to projected daily growth rates of 5.3-6.9% for Ontario and 6.6-10.4% for Quebec. Projections for daily deaths over the next two weeks also show higher levels of uncertainty in Quebec than other provinces, reflecting this province's higher past and projected future growth rates.

Finally, projections for the future number of deaths do not appear to be sensitive to assumptions regarding the infection fatality rate: increasing the fatality rate by 50% in our simulations does not alter the model's projections for deaths. However the higher fatality rate does affect the projected infections, reducing the projection for cumulative number of infected for Canada on May 2 by approximately 40%, from a mean of 594,000 to a mean of 361,000.

Table 2: Summary of the two-week forward projections for COVID-19 infections and deaths, as generated by the demonstration case study model for Canada and four of its provinces, and assuming current public health measures remain in place. Ranges represent the 95% Monte Carlo confidence intervals over 1000 Monte Carlo realizations.

Jurisdiction	Model projections for May 2, 2020	
	Cumulative Infections	Cumulative Deaths
Canada	398,000 – 2,023,000	3,370 – 4,250
Alberta	11,600 – 76,800	80 – 107
British Columbia	13,300 – 91,300	127 – 172
Ontario	141,000 – 826,000	1,170 – 1,460
Quebec	255,000 – 1,395,000	1,960 – 3,230

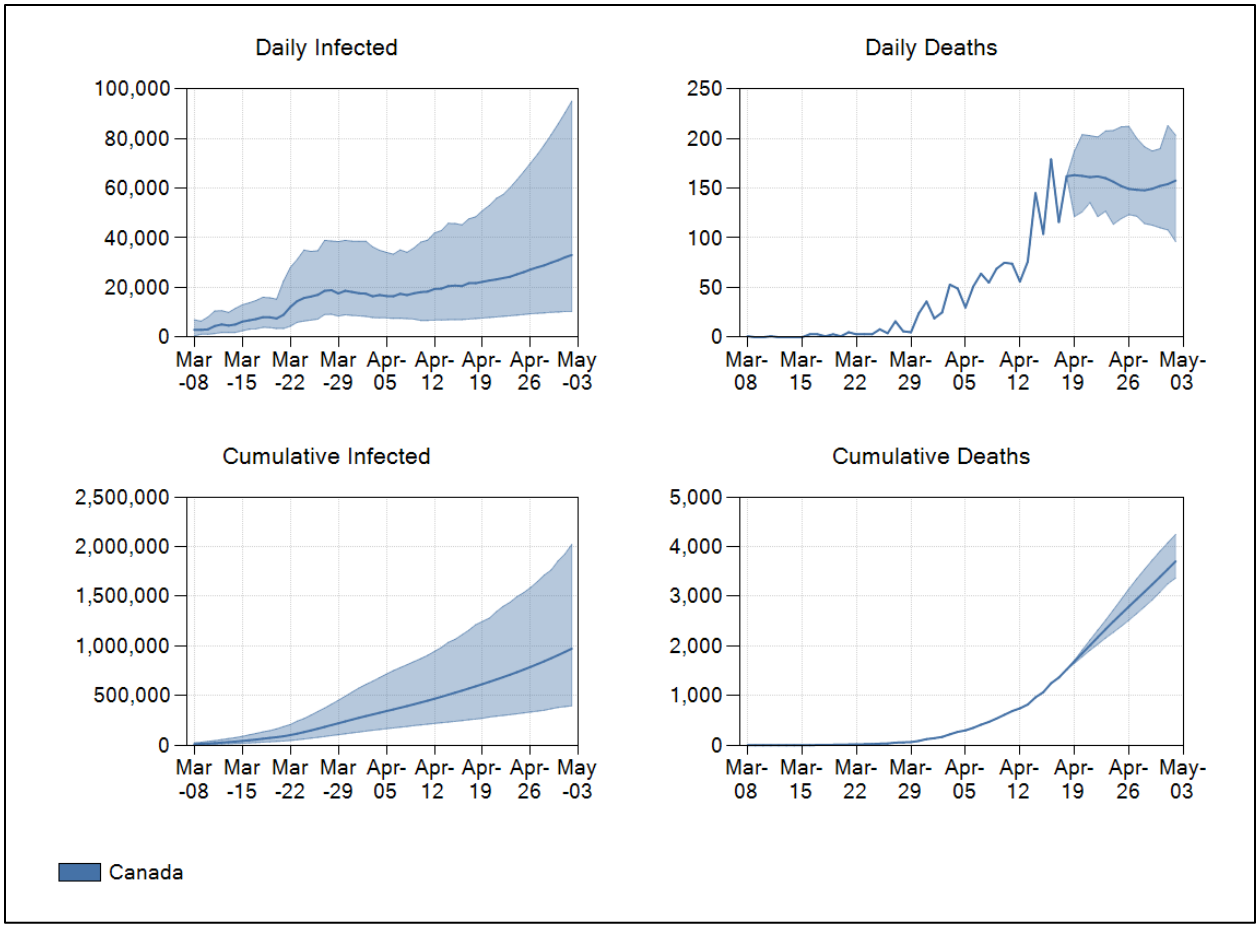

Figure 5: Model projections of COVID-19 infections and deaths for Canada, as displayed in the SyncroSim user interface. Solid line indicates the mean number of infections (daily and cumulative) and deaths (daily and cumulative) each day, while zones indicate the 95% Monte Carlo confidence interval over 1000 Monte Carlo realizations. Deaths up to and including April 18th are actual reported deaths.

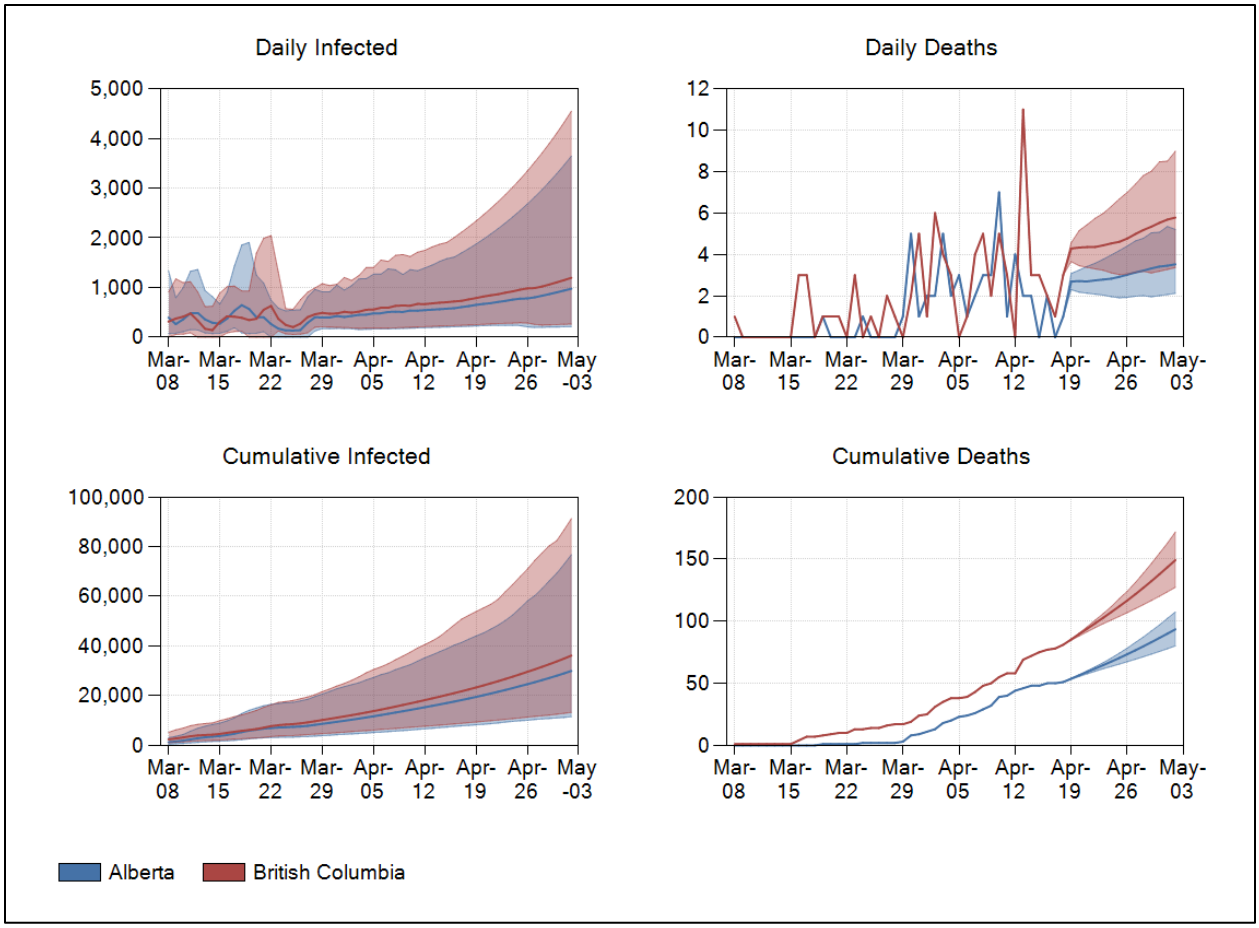

Figure 6: Model projections of COVID-19 infections and deaths for the provinces of Alberta and British Columbia, as displayed in the SyncroSim user interface. Solid lines indicate the mean number of infections (daily and cumulative) and deaths (daily and cumulative) each day, while zones indicate the 95% Monte Carlo confidence interval over 1000 Monte Carlo realizations. Deaths up to and including April 18th are actual reported deaths.

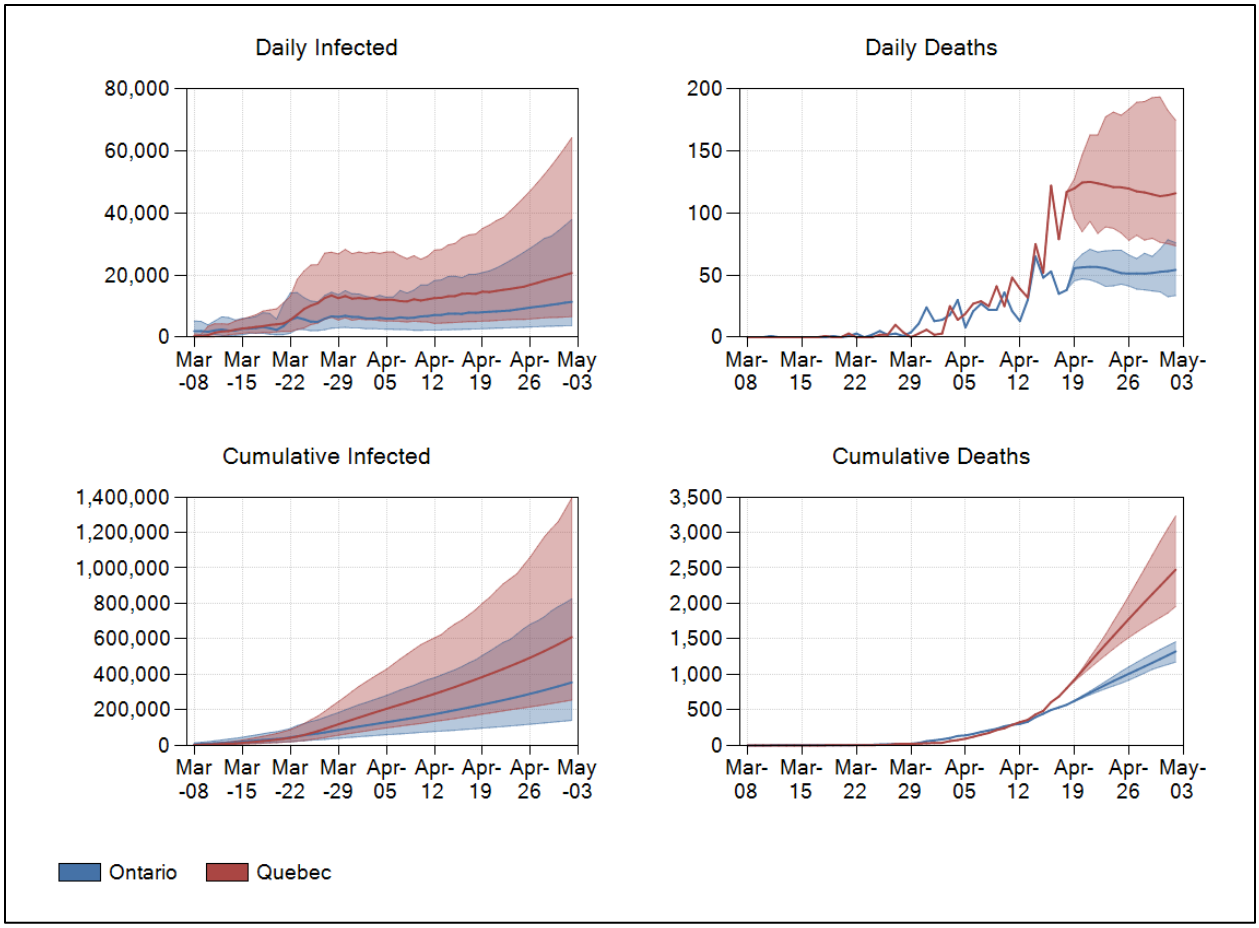

Figure 7: Model projections of COVID-19 infections and deaths for the provinces of Ontario and Quebec, as displayed in the SyncroSim user interface. Solid lines indicates the mean number of infections (daily and cumulative) and deaths (daily and cumulative) each day, while zones indicate the 95% Monte Carlo confidence interval over 1000 Monte Carlo realizations. Deaths up to and including April 18th are actual reported deaths.

Concluding Remarks

The SyncroSim modeling framework presented here provides an alternative approach to projecting COVID-19 infections and deaths. The framework can be configured for use with any jurisdiction, ranging from public health units to countries. The framework is stochastic, allowing forecasts to reflect uncertainties in model inputs. Finally, and most importantly, the framework is scenario-based, allowing users to develop and test their own jurisdiction-specific "what-if" scenarios for use with the model.

Our case study demonstrates how the framework can be configured to make short-term projections of infections and deaths for Canada and its provinces. While we considered here just two future scenarios – a scenario continuing current measures and one in which there were no public health measures – future efforts could easily consider a broader range of scenarios. For example once new death data comes online for reference jurisdictions that are currently beginning to relax public health measures (e.g. Spain), these new data could, in turn, be used within our existing Canadian model to forecast the

consequences of similar changes in Canada. A user friendly version of the model, pre-loaded with the Canadian case study data yet suitable for configuration for any jurisdiction, is available for download at www.modelthecurve.ca

Acknowledgments

The case study model presented here, along with the underlying infection and death model, was self-funded by Apex Resource Management Solutions. Core development of the SyncroSim software upon which this model was implemented has been supported by The Nature Conservancy, U.S. Geological Survey, U.S. Forest Service, the LANDFIRE program and Apex Resource Management Solutions. Special thanks to the many visionary scientists south of the 49th parallel, namely Mr. Ben Sleeter, Dr. James L. Smith, Dr. Louis Provencher, Dr. Zhi-Liang Zhu, Dr. Brian Miller, Dr. Catherine Jarnevich, Dr. Jennifer Costanza, Dr. Jeffrey Morisette , Dr. Wendel Hann and Dr. James Merzenich, without whom Syncrosim would not be here today.

References

Berry, I., J.-P. R. Soucy, A. Tuite, D. Fisman, and C.-C. O. D. W. Group. 2020. Open access epidemiological data and an interactive dashboard to monitor the COVID-19 outbreak in Canada. CMAJ. 2020 Apr 14;192(15):E420. DOI: https://doi.org/10.1503/cmaj.75262

Dong, E., H. Du, and L. Gardner. 2020. An interactive web-based dashboard to track COVID-19 in real time. The Lancet infectious diseases. DOI: https://doi.org/10.1016/S1473-3099(20)30120-1

Flaxman, S., S. Mishra, A. Gandy, H. Unwin, H. Coupland, T. Mellan, H. Zhu, T. Berah, J. Eaton, and P. Perez Guzman. 2020. Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. Imperial College London. DOI: https://doi.org/10.25561/77731

Lauer, S. A., K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler. 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Annals of internal medicine. DOI: https://doi.org/10.7326/M20-0504

Murray, C. J. 2020. Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months. medRxiv:2020.2003.2027.20043752. DOI: https://doi.org/10.1101/2020.03.27.20043752

Verity, R., L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. Walker, and H. Fu. 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet infectious diseases. DOI: https://doi.org/10.1016/S1473-3099(20)30243-7