Minnesota COVID-19 Modeling

FREQUENTLY ASKED QUESTIONS

UPDATED 05/13/2020

Version 3.0 Specific Questions – published 05/13/2020

When was Version 3.0 of the MN COVID-19 model released?

In the effort to model the course of the COVID-19 epidemic for Minnesota, the research team updates inputs and the structure of the model to best align it with available evidence and produce the closest possible reflection of real-world circumstances. Version 3.0 of the model was released publicly on May 13th.

What has changed in Version 3.0 of the model and why?

The updates to Version 3.0 reflects a growing understanding of COVID-19 and the availability of new data. Version 3.0 of the model includes structural changes, updated parameterestimates based on emerging U.S. data, and additional calibrated parameters from the model to fit the most recent data on MN hospitalizations and deaths.

As part of the structural changes, symptomatic and asymptomatic infections are now explicitly represented in the model. Ddeaths can occur outside of the ICU in the model. The ICU demand represents cases requiring ventilation (vs. just a critical care bed). Certain model parameters are now estimated by calibrating the model to cumulative deaths and prevalent hospitalizations from March 23 – April 25, 2020 (new to version 3.0). Another feature is that calibrated parameters include the percentage of contact reductions between people under different mitigation strategies, the percentage of infections that are mild or asymptomatic, and parameters used to estimate deaths occurring outside of the hospital.

Has the Minnesota model been formally evaluated and validated?

In preparation for the release of version 3.0, the research team has sought peer review from experts in academia and the federal government. The team is also preparing a manuscript for publication which will undergo a formal peer review as part of the publication process.

Why aren't all parameter estimates in the model from Minnesota or the U.S.?

Aggregate data on COVID-19 cases in the United States have begun to emerge; however, follow-up time is necessary to ascertain outcomes among active cases. Recent studies illustrate the potential extent of missing data.

Will Version 3.0 be the last update to the MN COVID-19 model? Why or why not?

No, the research team expects to make further changes over time. Though considerable refinements of the model have taken place over the past two months, the disease itself and what we know about the virus are still relatively new and incomplete.

At the same time, science is evolving extremely fast, and as critical questions are being answered over the next few months, the research team will aim to incorporate them into the model. Because much of the evidence is emerging in pre-publication settings to ensure quick dissemination. The downside of that approach is that the quality of the studies, the data and the conclusions has been mixed. This necessitates *careful* consideration of the evidence.

How can I access the Minnesota COVID-19 model code?

The Minnesota COVID-19 model code for the early version of the modle (not including the most recent 3.0 version) is publicly available online via Github (https://github.com/MN-COVID19-Model). Accessing the code requires a GitHub account (https://github.com/join), which is a free software development platform. Running the code requires R (https://www.r-project.org/), a free statistical software. The code runs base case scenarios described in the model slide deck. Generating uncertainty around base case scenarios is computationally intensive. As such, uncertainty analyses are not included with the code release.

Can someone assist me with interpreting technical details of the model code?

The programming code is available online (see previous question). Because the research team is responsible for activities related to continuously updating and refining the model, assistance with the code is not available. However, to the extent possible, MDH will continue to update this FAQ in response to general clarifying questions.

Why did the release of version 3.0 take a few weeks?

Typically, this science takes months, if not years. It involves developing a robust understanding of the dynamics of the disease, assessing the literature concerning key transmission characteristics, adapting the conceptual models to align with the pathogen, identifying and testing underlying data, writing programming code, and seeking iterative rounds of peer review.

Second Round of Questions (v.1 and v.2) – published 05/13/2020

COMPILED BETWEEN 04/10/2020 AND 05/12/2020

What are the differences between version 1.0 and version 2.0 of the model?

Version 2.0 of the model was fit to actual COVID-19 deaths in MN to inform the prevalence of undetected infection at the start of the model simulation. Uncertainty was incorporated in parameter estimates, and parameter estimates were updated with newly available data, where possible. A summary of updated estimates includes:

- Lower probability of hospitalization

- Higher R_0

- Shorter length of hospital stay

- Updated age and underlying condition distribution of model population

References for these estimates can be found in the model technical documentation available online: https://mn.gov/covid19/data/modeling/

Why did the value of R_0 increase in version 2.0 of the model?

Version 1.0 of the model used an R_0 value estimated from Wuhan data in China's Hubei province, which at that point represented the most complete available information. Version 2.0 of the model uses an R_0 value that leveraged newer data from 11 European counties with advanced epidemics. Data from these countries may be more representative of the epidemic in the U.S.

It's important to note that R_0 can vary by geography and population. Implementation of mitigation strategies also influence rate of infection growth and any empirical estimates of R_0 (or more precisely, the effective reproductive number, R_t). We anticipate continual updates to published R_0 values, particularly as more data from the US become available; results of this model will also change accordingly.

Why do results from the Minnesota COVID-19 model differ from other models? In particular, why are results so different from the IHME model that has been used by the federal government?

The conceptual approach, the assumptions, and the choice of parameter estimates will all affect outcomes produced by different models. In addition, models can differ in the period their forecasts cover. For example, while the Minnesota COVID-19 model describes the potential course of the epidemic over a full year, the IHME data generally restricts its forecast window to 4 months.

With regard to the model developed by the Institute for Health Metrics and Evaluation (IHME), there are considerable differences that are responsible for what are likely to be rather optimistic results for Minnesota (and other jurisdictions). The IHME model differs in a number important ways:

- Rather than being aligned to the epidemiologic course of the disease and what we know about transmission and fatality, the IHME model uses fatalities from the U.S. and states to fit to the mortality curve derived from Wuhan, China.

- Implicitly, this approach assumes forms and lengths of social isolation of infected persons that are not observed in the U.S. and perhaps not congruent with societal values in this country. The model also assumes the social distancing practices would stay in place as long as needed.

- The IHME model also projects outcomes just over a period of four months, as opposed to a full year or the full course of the epidemic, as done with the Minnesota COVID-19 model. Because of that it is inappropriate to make direct comparisons of the number of estimated COVID-19 fatalities derived from the two models – these would be apples to oranges comparison.

Are model estimates of infections and deaths restricted to patients who test positive (i.e., detected cases)?

The number of undetected infections in the population cannot be directly observed using case data, because not everyone who is infected develops symptoms, and not everyone in Minnesota receives a test. The model estimates the prevalence of undetected infections prior to the first day of the model simulation (March 23, 2020) to project infections and deaths that are inclusive of both detected and undetected cases of COVID-19. The percentage of infections that had been detected in Minnesota as of March 22, 2020 was calibrated by fitting the model to observed COVID-19 deaths. This estimate established the total number of prevalent infections in Minnesota at the beginning of the model simulation.

The exercise of fitting the MN COVID-19 model to deaths in the state resulted in an estimated case detection rate of 1 percent (version 2.0). That means that, for the early part of the epidemic, we assume the actual number of infections is about 100 times the number of detected infections.

Why do my "back of the envelope" calculations for infections and deaths not match the model estimates?

As described in the previous question, the model estimate for the proportion of infections detected prior to March 23 should only be used when establishing the initial conditions of the model. This estimate should not be applied to current case data to calculate the prevalence of all infections in Minnesota, as detection rates vary over time due to testing and surveillance practices. Additionally, it is important to properly interpret the model parameters when attempting to do "back of the envelope" calculations. For example, in version 2.0 of the model, the probability of hospitalization is limited to symptomatic cases, and the mortality rate is restricted to cases in the ICU. The design of these parameters, which are age-dependent, is

based on data available for detected cases only and the assumption that undetected cases are less severe. Although the model is able to estimate the actual number of prevalent infections (i.e., both detected and undetected cases), the risks for hospitalizations, ICU need, and death cannot be uniformly applied across all infections.

Why do some data in the Minnesota COVID-19 model come from China?

The data we are using to inform the model are continuously changing as our knowledge about this pathogen expands. With evidence about COVID-19 slowly emerging from across the globe, reports from early outbreaks in China provide the most complete clinical and epidemiologic data available. This is beginning to change, with evidence emerging from European and U.S. settings, albeit very slowly. Wherever possible, we are using US and Minnesota-specific data.

Does the model capture economic harm associated with mitigation strategies?

This model is an epidemiologic model that was designed to simulate disease transmission. The model was not designed to forecast economic outcomes. That said, it is clear that the pathogen and the mitigation strategies to delay and reduce its impact affect Minnesotans' health and economic wellbeing.

It is important to remember that model output represent just one of many sources of information being considered by state leaders who shape our shared response to COVID-19. The views of stakeholders, evidence from the economy, public health perspectives, changes in health care seeking behavior, and other factors affect decision-making.

Why are the impacts of treatments or a vaccine not being modeled?

Models can only forecast the impact that external factors might have on disease transmission if key aspects of their functioning are understood. For example, what proportion of the population is impacted, is the disease being prevented or only its severity being reduced, how long would it take for an effect to materialize?

While the development of treatments and vaccines are underway, preliminary effectiveness results have been reported for only one treatment (Remdesivir) thus far. This evidence had not emerged until after v.2 modeling was complete. We have modeled some scenarios using version 3.0 of the model of the potential impact of treatments like Remdesivir, a Gilead antiviral drug, on COVID-19 outcomes, though it is important to emphasize that the clinical trial evaluating this treatment is still ongoing.

Can you run the model on different populations (for example, a specific Minnesota county)?

Uncertainty in model results increases with uncertainty in parameter estimates, which can result from limited data and small sample sizes. The MN COVID-19 model is calibrated to data from the whole Minnesota population, yet the outcomes are still associated with significant uncertainties. Minnesota is still relatively early in the epidemic with limited data available.

A COVID-19 model developed for jurisdictions *smaller* than the whole state would be associated with more uncertainties, possibly to the extent that it would not add meaningful information or may create a false sense of precision. More importantly, in order to simulate disease progression in different populations, a model would require the availability of data specific to that population - contact patterns, demographics, underlying conditions, and the volume of COVID-19 cases and fatalities. Future iterations of the model may explore possible extensions of the model to different Minnesota geographies depending on the availability of these types of data.

How is uncertainty incorporated in model results?

Recognizing that mathematical models are associated with considerable uncertainty is important for their interpretation and use. Model outputs typically represent general directions and scale, rather than precise dates or volumes.

Uncertainty can arise from a number sources, including the completeness of available data and the choice of parameter values. The MN COVID-19 model employs many parameters, and these parameters require estimated values. For example, the MN COVID-19 model has a parameter for how much more likely it is for individuals with a comorbidity to die COVID-19 after they have been hospitalized (compared to others without a comorbidity. We used evidence in the literature to establish the base value of this parameter for the model. However, there are a range of values around this base estimate that are also plausible. In a process known as sensitivity analysis, the impact of uncertainty in parameter values is explored by running repeated simulations of the model with different parameter values sampled from their plausible ranges. Each run of the model with a different set of parameters produces a different set of results. Collectively, these results represent possible outcomes accounting for uncertainty in the parameter values. An example of results with uncertainty can be viewed in "spaghetti plots", where output from each simulation is represented with a distinct curve, in the model slide deck (https://mn.gov/covid19/data/modeling/). Presenting the results generated from each parameter set as its own curve (as opposed to reporting as an average) is necessary to preserve the correlation between the outcomes at each time step.

What underlying conditions are associated with an increased risk of COVID-19 hospitalization and death?

Over the past few weeks, new evidence is emerging about what underlying health conditions seem to be associated with worse COVID-19 outcomes. For version 2.0, these underlying conditions are hypertension, diabetes, cardiovascular disease (ischemic heart disease or heart failure), chronic obstructive pulmonary disease (COPD), chronic kidney disease, and cancer. This list was compiled by reviewing the literature to date and will be refined and updated in futures versions of the model.

In version 3.0 of the model, an increased risk of death due to underlying conditions is not explicitly modeled due to limited complete U.S. data. In future iterations, as more robust data become available on specific comorbid conditions and the increased risk of death in the ICU, hospital or other settings become available, we will update our model accordingly.

What exactly does the unit for ICU mortality – per 10 person days – mean? This is something I saw in the v.2 material.

The concept of "person-time" is used in the field of epidemiology to incorporate the amount of time someone is at risk of a certain outcome in the denominator of measures. It is a way to standardize across populations with different experiences. Mortality is often expressed in terms of number of deaths divided by the person-time over which a group of individuals is observed. For example, 5 people each observed for 2 days would be an example of 10 person days, as would 1 person observed for 10 days. Another way to say *per 10 person days* is *per 10 persons per day*. The following link is to a CDC introduction to some common epidemiologic measures, including person-time: https://www.cdc.gov/csels/dsepd/ss1978/lesson3/section2.html

How many Minnesotans are expected to get SARS-CoV-2 in the model?

The purpose of the Minnesota COVID-19 model is to estimate (over the predictable course of the epidemic) how many people are likely to contract the pathogen responsible for COVID-19, require health care services, and die, as well as when that is likely to occur. As such, it represents a simplified representation of the world.

COVID-19 stems from a virus that is new to humans, which means that few people have immunity against it and, thus, virtually everyone is at risk of becoming infected. The Minnesota model estimates that, over a full year, up to 88% of the Minnesota population, or 5,005,000 people, could become infected. This includes people who have symptoms, and people who do not. Generally, transmission of the virus is expected to take place until sufficient immunity (sometimes referred to "herd immunity") has been established in the population, either naturally through ongoing transmission over time or through the availability of an effective vaccine. In addition, it is important to recognize that these estimates can vary with mitigation strategies in place.

First Round of Questions - published 04/10/2020
COMPILED PRIOR TO 04/10/2020

What is the Minnesota COVID-19 Model?

An interdisciplinary team at the University of Minnesota has been working with the Minnesota Department of Health to use available evidence on COVID-19 to estimate the trajectory of the disease in Minnesota using a SEIR (susceptible, exposed, infected, recovered) framework. This model aims to understand how the COIVD-19 epidemic will evolve in Minnesota and how

different social distancing policies may impact it. The model accounts for state-specific demographics and the prevalence of underlying health conditions, and is calibrated to recently observed Minnesota COVID-19 mortality data. The model was developed to help inform the state's policy and operational responses. Details on the structure of the model, assumptions, and the underlying data are available at www.mn.gov/covid19.

What does the model tell us and what does it not tell us?

Mathematical models are not crystal balls. They cannot tell us what will happen in the future. Instead they provide a range of plausible outcomes given what is presently known about disease natural history and the state of the epidemic. Given the novelty of this virus, our understanding of these things is incomplete. Model outputs are therefore associated with considerable uncertainty. These uncertainties can be estimated and included in presentations of model results. The Minnesota model uses data from other states and countries where community spread started earlier to project estimates of daily COVID-19 case counts, intensive care unit (ICU) bed occupancy, and deaths in Minnesota. The model tells us *approximately* when the peak of the epidemic is expected to occur, what the expected magnitude of the peak is, and when the required number of ICU beds could be expected to exceed the state's capacity.

What have we learned from the model so far?

Select results from the Minnesota model have become available from the first two versions of the model. At a high level, the model results show that social distancing strategies similar to what has been recommended are expected to postpone the timing of the peak in infections and ICU admissions due to COVID-19.

Why did we develop the Minnesota COVID-19 Model?

When new diseases emerge, policy-makers, business leaders, and public health officials must make decisions under conditions of uncertainty. The newness and severity of this disease means that decision makers do not yet know the best way to control this epidemic. By combining available data on how the virus causes disease (what epidemiologists call "natural history") with mathematical equations, researchers can create models to forecast key outcomes such as the number of critical cases or deaths. The ability to forecast will, however, depend on the quality and availability of data. For a new disease like COVID-19, much remains unknown or uncertain. To address this, models can produce results that reflect this uncertainty.

How are state officials using this model?

The Minnesota Department of Health is using the model to simulate different hypothetical social distancing measures to predict their impact on the timing and reduction of COVID-19 cases and deaths in the state. Modeling the number of people who have the disease each week helps to establish the approximate window of time from when mitigation strategies were implemented until the point at which the state's health care systems might become overloaded by COVID-19 cases, assuming current levels of resource availability. Governor Walz and other

state leaders use results from this model to inform response strategies, such as the stay-at-home order, and plans for increased health care capacity.

Why don't we see a flattening of the curve with mitigation strategies?

Social distancing is an effective tool for reducing transmission but it cannot stop it. Flattening the curve would require *prolonged* social distancing over the course of many months. Currently the model compares relatively short social distancing measures. For this reason, differences across modeled scenarios appear as a delay in the epidemic peak but not a flattening of it.

Where did model parameter estimates come from?

In a model, parameters are the values that reflect assumptions about the speed and probability with which people progress through various states (from susceptible to infected, for example). Values for this model are informed by published reports using data from early outbreaks in China, Europe, and the U.S. They are also carefully vetted to consider data quality and relevance to Minnesota. Where multiple plausible estimates are available, they inform the range of uncertainty around that estimate. These studies represent the most complete data to date on the clinical characteristics, risk factors, and outcomes of COVID-19 cases. A full list of model parameters with sources can be found in the model technical documentation available at www.mn.gov/covid19.

It is a common practice to update models as new evidence becomes available to ensure model inputs are drawn from the most up-to-date information on similar populations. The second version of Minnesota's COVID-19 model updated the following parameters in light of newly available data: the transmissibility of the virus (higher), the average length of COVID-19 related hospitalizations (lower), and the proportion of infected individuals requiring hospitalization (lower).

How are data on Minnesota's confirmed cases used in the model?

Information about confirmed cases in Minnesota are used to calibrate the model to ensure that it reproduces case counts similar to what was reported in the first 20 days of the state's epidemic. The estimate of the percentage of infections that are detected (confirmed cases) was adjusted until the model reproduced daily death counts that were similar to actual death counts. This process established an estimate for the prevalence of undetected infections on the first day of the model simulation – a parameter that cannot be directly observed. It increased estimates of the *total* number of infections in the state by a factor of approximately 100.

Why do results from the Minnesota COVID-19 model differ from other models?

The conceptual approach, assumptions, and parameter estimates will all affect outcomes of each particular model. In addition, models can differ in timescales. Unlike other shorter-term models, the Minnesota model describes the potential course of the epidemic for a full year.

Divergent results from different models is not unexpected. However it is important to clarify the particular model features that explain these differences. The Institute for Health Metrics and Evaluation (IHME) model, for example, predicts far fewer cases and deaths for Minnesota. This is due to several reasons. First, the IHME model assumes far more restrictive social distancing measures implemented for significantly longer periods of time. It also projects outcomes over the course of four months as opposed to a full year as with the Minnesota model. Finally, it does not explicitly account for the elevated risk of illness and death from COVID-19 associated with underlying health conditions.

MDH has referred to the second version or iteration of the Minnesota model. Will there be future versions, and if so, what will be different?

It is common practice to refresh models when new data become available. As with the second version, subsequent versions will continue to incorporate the most up-to-date information about COVID-19 in Minnesota, the U.S., and across the globe. We will adjust model parameters appropriately based on new data. Updates will impact model output and ultimately reduce uncertainty.

Moving forward, the model structure and output may also be modified to better meet the needs of policy makers and other decision-makers. For example, we may add features to allow users to compare the timing of certain mitigation strategies, to account for geographic variation, and to incorporate the potential effects of enhanced access to antibody testing.

What do the ranges in the model results mean?

The ranges are uncertainty intervals around model outcomes. The intervals produced by this model give a range of results that might be expected given the uncertainty around parameter values used.

For more information, contact:

University of Minnesota School of Public Health
Eva Enns (eenns@umn.edu), Shalini Kulasingam (Kulas016@umn.edu)
Media: unews@umn.edu

Minnesota Department of Health
Stefan Gildemeister (Stefan.Gildemeister@state.mn.us)
Media: health.media@state.mn.us