Article details: 2020-0093	
Title	Estimated surge in hospitalization and intensive care due to the novel coronavirus pandemic in the Greater Toronto Area, Canada: a mathematical modeling study with application at two local area hospitals
Authors	Sharmistha Mishra MD PhD, Linwei Wang MSc, Huiting Ma MSc, Kristy C.Y. Yiu MSc, J. Michael Paterson MSc, Eliane Kim MPH, Michael J. Schull MD MSc, Victoria Pequegnat MSc, Anthea Lee MMA, Lisa Ishiguro MSc, Eric Coomes MD, Adrienne Chan MD MPH, Mark Downing MD, David Landsman BSc, Sharon Straus MD MSc, Matthew Muller MD PhD
Reviewer 1	Dr. Dena Schanzer
Institution	Public Health Agency of Canada, Infectious Disease and Emergency Preparedness Branch, Ottawa, Ont.
Reviewer comments and author's response	GENERAL (ABSTRACT, STUDY OBJECTIVE) 1. My first impression from the abstract is that the scenarios are already out of date. This is a major challenge in publishing forecasts during an evolving emergency. The abstract should contain enough information to give the reader an appreciation of what they would learn by reading your paper. While the objective of the project was to provide the two hospitals with an estimate of potential resource requirements during the early stages of the shelter in place orders, the objective of publishing these scenarios is not obvious. I'd suggest rewriting the abstract with a focus on your audience and include the objective for publishing the model. I have provided a few thoughts on this basic question. Response: Thank you for the helpful suggestion. We have revised the paper as suggested by the reviewer and editorial team to take a historical perspective. We have accordingly revised the abstract with the objective. 2. As we are still in the middle of an ongoing epidemic, this work would likely still be very useful in preparation for a fall wave, or the predicted second wave once the emergency restrictions or lockdown measures are eased. If you are continuing to advise the hospitals on resource requirements over the next couple of months, it would seem that the model would need a mechanism to respond to local epidemic growth rates. I am thinking of how this is handled in other emergency situations such as a hurricane? What are your plans to update this model as the situation evolves? Or to select the appropriate scenario to better match the current situation? What data would be needed to update your model over time? Response: We appreciate the excellent questions, and one option as suggested by the editors was to consider framing the paper as model description paper – for current and future use as a flexible tool. As a team, we felt that it was most appropriate to re-frame the paper from a historical perspective. We do have plans to update the model for additional heterogeneity and future use, and will be presenting and submitting the new model as a separate study and future work. 3. The epidemic growth rate has been unusually variable over time in many countries – slowing considerably once emergency shelter in place orders were declared. Unfortunately, the epidemic growth rate in Ontario (and Canada) was slow to respond, perhaps responding to the super-spreading events of the return of a very large number of Canadians through over-crowded airports in mid-March, and then the spread in long term care facilities two weeks later. We were also slow to increase our testing capacity, which again would impact the rate of increase in the number of confirmed cases. Quite the challenge to find timely data to inform

	these prediction models. Response: At the time of the model development and analyses which were used to inform the hospital response, there were limited data. However, these data are now available and will be used in future studies. 4. I would think it is a given that ICU beds are insufficient to handle the surge. How have other hospitals handled this? What decisions were made by the hospitals? How did they use your model results? Response: We cite approaches taken from SARS in 2003 [1]and recommendations that followed (which were a guiding principle for this analyses in March 2020), and outline in the discussion how the two hospitals ceased elective surgeries and also re-organized staffing and back-up plans to ensure both bed and staffing capacity may be able to adjust. (Discussion, Page 6): “The preliminary hospital-specific findings (on March 4, 2020) were used to prepare for the local surge at the two hospitals, with the updated analyses from March 25, 2020 used to continue planning efforts. First, the hospitals opened up beds by temporarily cancelling non-essential surgeries and procedures. Second, as most COVID-related inpatient care would fall under the hospitalist and medicine services, the relevant departments rapidly set up a separate service with a viable back-up system and ability for rapid scale up in anticipation of increasing cases requiring admission, and staffing short-falls due to infection, exposure, or while awaiting test results if symptomatic. Third, ambulatory clinics were reduced with a focus on virtual care and urgent assessments only; this allowed clinic space to be consolidated to preserve personal protective equipment and human resources (including physicians) for deployment to other areas. This consolidation also allowed identification of potential inpatient spaces. There was also a change in health-care use by the public: non-COVID medicine admissions are dropping across the city and country (29). The active and passive reductions in admissions meant that neither hospital went over-capacity." [1. Schull MJ, Stukel TA, Vermeulen MJ, Guttmann A, Zwarenstein M. Surge capacity associated with restrictions on nonurgent hospital utilization and expected admissions during an influenza pandemic: lessons from the Toronto severe acute respiratory syndrome outbreak. Acad Emerg Med. 2006;13(11):1228-31.] 5. As it has become evident that the exponential growth rate has responded to the various levels of lock down, doubling time has become the main communication tool to describe the rate of spread. R0 is not sufficient to characterize an epidemic, as the generation time, which to some extent is also unknown, is also needed to calculate doubling time. As well, the scenarios (fast, default, and slow) are not sufficiently described in the abstract to be meaningful. Figure 2 in the main body of the manuscript gives the impression of a very large level of uncertainty, as if the scenarios (fast, default, and slow) were just arbitrarily picked. You mention that they are Interquartile Ranges based on a simulation with some fixed parameters? The results of two arbitrary scenarios are not very helpful. However, if you can succinctly describe a fixed dataset, for example, case data for countries meeting specific criteria as of a specific date as a basis for your scenarios, you could report on the IQ range or the range associated with percentile equivalent of a 95%CI for this dataset resulting from a bootstrap.

	Response: We agree, while some places use doubling time – the most common metric now is the time-dependent R (or Rt). We have clarified that the fast/default/slow scenarios were not arbitrarily chosen but selected based on pre-defined criteria within the constraints of scenarios from other settings. The use of scenario-based analyses was commonly used across countries to help guide planning [1] in the setting of an emerging outbreak for which little data were available at the time. We have also focused the re-framing of the paper (from the historical lens) on the default scenario. As epidemics are simulated and not fitted/calibrated, it would not be appropriate to provide 95% CI in this case. For the selection of scenarios, the following description is provided in the (Methods, Page 4): "We then selected a slow/small epidemic and a fast/large epidemic using the lower and upper interquartile range in the peak incidence across the full, constrained set of epidemic trajectories. We defined a default scenario using the median or best-justified parameter values which passed our internal validity checks and epidemic constraints." [1. Schull MJ, Stukel TA, Vermeulen MJ, Guttmann A, Zwarenstein M. Surge capacity associated with restrictions on nonurgent hospital utilization and expected admissions during an influenza pandemic: lessons from the Toronto severe acute respiratory syndrome outbreak. Acad Emerg Med. 2006;13(11):1228-31.] 6. Given that the March/April lockdown has been more successful than any of your scenarios, how best to present your work? For a medical audience, one suggestion is to describe the results for various doubling times, without specifically labelling scenarios which could quickly become misleading. (Use of doubling times for cumulative cases would not be appropriate for a second wave – best to use doubling times for new cases or hospitalizations.) Basing your results on doubling times could possibly have alerted the hospital administrators as to which scenario is more likely over the coming month. Response: Thank you and we approached the re-framing from a historical perspective, and provided dates accordingly. We clarified that the default scenario was used to inform the hospitals on March 25 (after an initial, preliminary analyses on March 4 prior to full review of data for model parameters). 7. Other assumptions also contribute to the uncertainty. For example, to estimate peak demand, we need a better idea of the severity of disease, and in particular, the proportion of infected persons whose symptoms are mild enough not to seek medical care. A sensitivity analysis which identifies the main drivers of uncertainty should be included and mentioned in the abstract. I'd also expect to see a tornado diagram in the body of the manuscript. Response: We provide a sensitivity analyses of all parameters included in the model. As we did not conduct model fitting, we did not conduct a tornado plot. In the absence of model fitting, the tornado plot would largely reflect wide or small input range. The dynamics are depicted in the panel of sensitivity analyses in the appendix. We included a statement in the abstract as follows: "Uncertainty in local epidemiological features was more influential than uncertainty in clinical severity." (Abstract: Results) 8. Perhaps results from serological studies to assess the potential level of

	exposure/immunity in the population will be available over the summer. If we have a better idea of how many people have been exposed to the virus, we would have a better idea of severity and could significantly narrow the range of uncertainty about the peak demand on hospital resources. Your discussion section should include a discussion of data requirements to improve the accuracy of your model. Given the large degree of uncertainty, a sentence on this should be included in the abstract. Given the huge costs associated with preparing and over-preparing or under-preparing for the peak demand, the need to be able to update any model on a timely basis as needed for preparation is urgent. Response: We have added to the limitation section as follows: “Finally, our objective was to conduct a scenario-based analyses, and not to explicitly fit the model to observed cases, hospitalizations, ICU admissions and deaths in the GTA. Future work involves adapting and fitting to population and setting-specific data on trends in testing, cases, and outcomes in the GTA.” (Limitations, Page 7) 9. Another possible option is to present the model from a historical perspective, looking back at the decisions made chronologically, and to some extent the lessons learnt. Again, dates are needed to follow your discussion and the situation is so fluid. Without this context, a number of your assumptions don't make sense. Response: This was a very helpful suggestion which we followed. We also included dates. INTRODUCTION 1. I would update the introduction after deciding on your objective for presenting the work. Response: We have updated the introduction based on the revised objectives. 2. The distinction between local and travel transmission is likely not that relevant at this point in time. Response: Given the historical perspective, we include mention of imported cases. However, we would suggest that all models continue to include imported cases because border closures are also dynamics, and re-opening of borders an important consideration. 3. If this model was used by the two hospitals to plan for peak demand, I'd suggest providing the context on this planning. As the situation was continually changing, it is very important to provide a date for the model. What was the cut-off date for model assumptions/parameters? When was the model presented to the two hospitals? [Editor’s note: this information should be included in the Methods section of the paper.] Response: We agree and have added the time-frames that were requested and used by the two hospitals pandemic command centers. We include the information in the Methods (Page 3): “The Unity Health Toronto COVID-19 Incident Management Team was formed on January 27, 2020 and requested rapid modeling in early March to estimate potential surge in health-care needs at each hospital. Preliminary results were provided March 4, 2020 and updated in late

	March using the constrained scenarios for a 6-week projection from March 25 to May 6, 2020." 4. I'd provide a short timeline: community transmission was identified on March 6, and this model was developed based on data as of March 10th?, leading to a decision to close schools etc. on … , and further restrictions to close all non-essential business on ..... Response: We provided a timeline of the use and development of the model, and its time-frame for use by the two hospitals. The model was used by two hospitals and thus informed the hospital specific response. METHODS 1. Why a model for just 2 hospitals? Response: The question came from the two area hospitals as that is the level at which decisions were being made (the province had yet to provide direct guidance on opening up number of beds, etc.) and that is also where the action to address surge were being implemented with immediate needs for surge planning. The questions were framed as recommended from the experience with SARS in 2003 [1]. [1. Schull MJ, Stukel TA, Vermeulen MJ, Guttmann A, Zwarenstein M. Surge capacity associated with restrictions on nonurgent hospital utilization and expected admissions during an influenza pandemic: lessons from the Toronto severe acute respiratory syndrome outbreak. Acad Emerg Med. 2006;13(11):1228-31.] 2. When was this model developed, i.e., why are your scenarios out of date? Response: The model was developed in March with the latest update on March 25, 2020. As such all data used to inform the model and the scenarios generated were based on the available knowledge at that time. 3. How will this model be updated, because we are not over this yet? Response: The model has been updated and revised considerably to include heterogeneity in transmission across outbreak settings and by age in the Greater Toronto Area, but will comprise a future study and submission. 4. I'd suggest moving the study setting into the intro. [Editor's note: please leave in Methods section] Response: We have retained the study setting in the Methods. 5. Doubling time after the 1st 3 cases would include an odd period, one that likely evolved from initially fairly restricted testing, (travel related only for Italy and then opening up to severe symptom based) before a major effort was made to increase testing. Looking back from what we now know, this period would likely over estimate the exponential growth rate. (if this is a limitation due to timing, then it should be mentioned in the discussion section. I don't think we initially realized that the early estimates of the exponential growth would soon start to decline, as they did in many countries even before the lock-down.) Response: We have included a time-variant testing strategy to reflect what

	happened with testing recommendations and criteria in the province of Ontario (Appendix 1, section titled “Modeled estimate of confirmed cases”). However, as we did not explicitly calibrate the model, nor compare the model to test positivity (data on test positivity were not available at the time of model development and analyses), we agree that using cases alone is a major limitation. We note this in particular given the different parameters for hospitalization/severity, which leads to more detection, in the Results Section (Page 5): “However, cumulative confirmed cases (Figure 2, Appendix 2 Figure 2.1) were much lower in the default scenario because of the clinical parameters: the proportion of individuals with COVID-19 with severe disease requiring hospitalization, and thus, detected, was 10.4% in the fast/large vs. 5.5% in the default scenario.” 6. Many of your figures (simulations) use days since the outbreak started. When did the outbreak start? It would help to be able to link any of these simulations to actual data or actual events. Response: We assumed the outbreak had been seeded on February 23, 2020. We include dates on the main figures for the hospital surges, and dates for the historical analyses that informed the hospitals’ pandemic planning. RESULTS 1. I find the results section difficult to follow and I don’t get a lot out of the results section. Response: We have revised the results with the historical perspective and hope it has improved the results section as well. 2. Given the large amount of uncertainty, a good summary of the sensitivity analysis would be helpful. It might be helpful to select a few outcomes of most concern to the hospital administrators, such as the timing of peak, and peak demand, and perhaps monthly demand. Tornado diagrams for each of the main outcomes would be easier to follow than text. If your range is an IQ range this should be clearly labeled, as the 95%CI percentile equivalent is more often used for credibility ranges. Response: We provide a summary of the sensitivity analyses, keeping in mind the word count and the objectives. We include all sensitivity analyses figures in Appendix 2, and outline the main source of uncertainty in the Results (Page 6, Section heading “Sensitivity analyses for hospital-specific surge estimates”). Given that the epidemics reflect a scenario-based analyses and not a fitted set of epidemic realizations, it would not be appropriate the report confidence or credible intervals. 3. Perhaps you could incorporate the app in some of the discussion. It was helpful to look at your app, though I could not easily understand why your estimates were so high, at least compared to the current situation as presented in the media. Response: We have reframed the paper from a historical perspective, and discuss the challenges with estimating/projecting beyond a short time-frame. Indeed, the longer-term projections are to reflect scenarios where nothing additional is done –

	and thus, to inform what could be done to reduce that trajectory. The observed epidemic and the projected epidemics should always differ beyond the short-term: because otherwise it means that we maintained the status quo and did not intervene. The Shiny R application was developed to allow individuals to replicate the analyses and interrogate the parameters and assumptions directly. Due to word count and prioritizing revisions to address the other suggestions (in particular, to allow for more room for the limitation section), we were not able to discuss the application in more detail in the Discussion section. 4. I noticed that your model derived epi curves were bell shaped or at least symmetrical. Most epi curves (daily # of new cases) are now showing a long plateau period followed by a slow decline. Even the media is remarking on this. Response: The epidemic curve follows that pattern based on the interventions remaining at a stable-level. The analyses we conducted for the hospitals assumed stable intervention levels, but the sensitivity analyses were conducted to understand the role that community-level interventions could play in modifying surge. What transpired after the early phase of the outbreak reflected various intervention strategies. INTERPRETATION 1. Dates are needed throughout the discussion in order to follow the discussion. It sounds like the hospitals started to prepare for a large surge around March 4th, i.e., before your model was developed. Response: We have added the dates to the figures, and in the main text (Methods, Results sections). 2. The results of the sensitivity analysis or analyses should be included in the discussion section, as data limitations seem to be the main limitation. [Editor's note: as you have done, the results need to be presented first in the Results section before being discussed in Interpretation, including the Limitations subsection.] Response: We have included in the Discussion section (Discussion, Page 6): "The local transmission dynamics, or what was happening with the epidemic overall in the city with respect to community-level interventions for example, had a larger influence on projected hospital's surge than uncertainty around disease severity." 3. Your scenario-based approach may still work. You may be able to find a better way to present this approach. Presenting 95% credibility ranges from your simulation may be more appropriate for the main body, perhaps leaving specific scenarios to the appendix, or accessible via the online app. I'd suggest providing some guidance as to when a particular scenario would be relevant. Response: We have outlined the rationale for selecting a default scenario as we did not conduct model fitting (Methods, Page 4): "To generate a plausible range of epidemic trajectories under best and worst-case scenarios, we sampled parameters as per Table 1 while fixing the intervention parameters, and used the following constraints: the upper and lower bound of the per-capita, cumulative cases detected per day in Lombardy, Italy, and Hong Kong, China, respectively,

	within the first 30 days after detection of 3 cases. We then selected a slow/small epidemic and a fast/large epidemic using the lower and upper interquartile range in the peak incidence across the full, constrained set of epidemic trajectories. We defined a default scenario using the median or best-justified parameter values which passed our internal validity checks and epidemic constraints. We also examined the face validity of our default epidemic by comparing it to our synthesis the observed data in the GTA where the first 3 cases had been detected by February 23, 2020 (Appendix 1)."
Reviewer 2	Dr. Jianhong Wu
Institution	York University, MITACS Centre for Disease Modeling, Toronto, Ont.
Reviewer comments and author's response	The authors are congratulated for a fine piece of study with anticipated influence on decision making and logistic preparation. Response: Thank you for positive feedback. 1. The authors have made it explicit that the transmission dynamics model is assumed homogeneous mixing. As COVID-19 case severity is age-specific, the authors may want to reiterate in the study limitation the need of refined models to take consideration of heterogeneous mixing in the population. Response: We have included in the Limitations section (Limitations, Page 7): "Future work includes capturing heterogeneity within the five health units and near real-time adjustment of the catchment using observed patterns of hospital-specific admissions; and heterogeneity in contact patterns by age-group and across congregate settings."