SEIR COVID-19 Compartment Model Overview

Margaret Harris Brockman: Welcome everyone to this webinar, SEIR Covid-19 compartmental model, a model to gain insight into the effect of interventions on outcomes. Thank you for joining us today. My name is Margaret Harris Brockman and I am with the National Collaborating Centre for Infectious Diseases. I will be your moderator during this hour. NCCID is funded by the Public Health Agency of Canada to provide knowledge and evidence for us in Public Health planning and [health] policy. Before we start, I would like to acknowledge that NCCID is located on original lands of the Anishinabe, Cree, Oji-Cree, Dakota and Dené peoples and on the homeland of the Métis Nation. We respect the treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in the spirit of reconciliation and collaboration.

Today's session provides an exploration of the SEIR model developed by modellers at the Public Health Agency of Canada. NCCID and PHAC are partnering to share models and supporting information for users across Canada. You can see all of the related resources at NCCID.ca and there are additional materials in development. So now, I'd like to introduce today's speaker. Aamir Fazil is Chief of the Risk Integration Synthesis and Knowledge Section at the Public Health Agency of Canada. Mr. Fazil holds engineering degrees from Drexel University and has been working in quantitative microbial risk assessment for more than 20 years. This includes being a primary contributor to several Work Health Organization food and agricultural organization microbial risk assessment simulation models.

Together with other scientists at the agency, he contributed to the development of the Covid-19 SEIR model. In today's presentation, Aamir will explain how the model works, and how you can adapt it to your local needs. We'll run the presentation and then he will be available live to answer your questions.

Aamir Fazil: Hello everyone. Today, I'm going to be talking to you about an SEIR model that we developed at PHAC to provide some initial insights into the Covid pandemic. I'm going to be describing is one of the models developed at PHAC. We've got a suit of models that have been developed and those will

also be shared on the NCCID site over the coming weeks. This particular model is a very traditional compartmental model. It was developed in Analytica, primarily because Analytica is a tool that allows you to do rapid development of models, it's also relatively graphic in the way its models are structured. They use inference diagrams to draw relationships between parameters. It allows model construction to proceed in a relatively insightful way, especially when you're dealing with both a technical audience, and perhaps subject matter experts that can provide some back and forth communication.

The focus of this model was to gain additional insights into what was happening, get some advice in terms of what might need for the refinement and hopefully inform some of the additional model development that we would proceed with. So, the overall objective of this webinar will be just to give you an overview. I'll, after this series of slides, walk you through the model itself in Analytica. I'll show you how to use the model, where some of the inputs and outputs can be accessed from, and how to understand the model itself. Then, I'll show you some of the results and some of the nuances around interpreting those results so that, hopefully, you'll be able to have a better understanding of the model, and maybe even modify and apply the model as appropriate for your situations.

So, the overall structure of the model is, as I said, a compartmental model and a very traditional SEIR model. We're looking at susceptible, exposed, infectious and recovered/removed populations within this model. That's shown on this slide on your screen now, where you have the population that's susceptible, potentially getting exposed to the infectious population. They then moved into an exposed category. The exposed people eventually become infectious themselves. When they are infectious, they can either be symptomatic and severe, symptomatic and mildly ill, or asymptomatic. When the model was constructed, there was a large degree of uncertainty at the time with what proportions were in what compartments.

And, if you've kept up with the literature, even today, there's still ongoing uncertainty as to what proportion of the population is actually asymptomatic, what proportion is mild, what proportion is severe, what the relationships are between those components, and potentially different demographics in the population. So, while we've learned a

lot, we're still learning, and that, in essence, is the story of modelling COVID in this epidemic across the globe. It's the constant generation of new knowledge, the constant refinement of models, and the need to always look at the new evidence, and apply that evidence to the models. Because, what you constructed a week ago can, many times, become outdated.

You need to constantly evaluate your parameters, the values, etcetera. No values you see that we might discuss in this webinar should be taken as fact and accurate for eternity. Things will change. You should scrutinize any of these values given what is known at the time you're going to be applying the model. What we have done within the model is used what were considered to be best estimates at a certain point in time. In Analytica, you can apply uncertainty distributions as well in your model. I'll demonstrate - or at least I'll show you where you can do some of that. That is good practice for the most part. You want to capture uncertainty in your parameters.

Overall, this picture shows you what the general structure of the model is. This slide shows you what you will potentially see when you open up your Analytica model. So, having given you that overview, I'll switch over to Analytica now, and walk you through some of the modelling process.

This is Analytica, the model opened up in Analytica, and this is typically what you would see when you open up the program. This is your input screen, if you will. It's broken up into a couple of sections that are probably appropriate to explain. You have your input parameter area. This is where you would enter the input parameters or what you know about either the population or the pathogen of interest.

You've got things like contact rates, how often people are, on average, contacting each other, the latent period of the infectious - of the infection, the infectious period of the pathogen, the symptomatic proportions, proportion severe, and time until hospitalized. The blue boxes represent some crude or general interventions, non-pharmaceutical interventions, or NPIs. These would represent Public Health measures that we may implement, that we might be interested in looking at to see what effect different strategies might have on the general epidemic trajectory. There is the results section, which by clicking one of these buttons will generate appropriate results, and from there, you can

explore the interventions, the results, general curves, and get some insight from them.

Next, the large blue box here gives you access to the model itself, the logic in the model, how the equations are structured, and how parameters interact with each other. Just to give you an example of that, I'll open up the model structure. When you do that, this is where you see, if you will, the real machinery behind the model - the influence diagram, which I mentioned earlier on. Analytica is constructed primarily using influence diagrams. Inference diagrams are, although you might be tempted to think of them as a flow chart, not a flow chart. They represent how parameters influence other parameters. So, the arrows don't necessarily represent a flow of - either the flow of people or flow of things. Rather, they show how one parameter influences another parameter.

In this particular case, you can see the delta susceptible, a change in the susceptible population, is influenced by the transmissibility with contact, or the probability of infection when someone comes in contact with an infectious person. That's how you should read the diagram. A bubble and an arrow pointing to another bubble tells you that that bubble has an influence on the receiving bubble. So, what you find when you open up one of the bubbles is a description of what that particular parameter is. In this case, as I said, delta - the variable name is DE. There is a description of what is happening in this particular thing – the title. In here, we see it estimates a rate of which exposed people move into the infectious category. Then, it gives you the equation as to how that parameter is influenced.

You can see it's one over the latent period times the exposure to severe parameter. All that to say, Analytica is a very graphical display of how the model works. You can quickly see how a parameter influences another parameter. If you open up that parameter, you'll get the additional details of what's happening behind that value. That applies for everything within the model. It would allow you to see every parameter, what the implications are of that parameter, how it influences other parts of the model, and which parts of the model influence it. Having said that, we'll spend a little time exploring some of the results to show how things can work. I didn't mention that - this is where you have your input parameters and Analytica does allow you to use distributions of values that could represent your uncertainty.

Right here, we've used a contact rate of 10. That means that each person is having, on average, 10 contacts a day prior to any social distancing interventions applied. In theory, you could apply a distribution in here. So, for example, you could use a normal distribution, a uniform distribution, or any distribution that you might think represents the uncertainty. You could enter a uniform distribution - say, you think that it should be uniform anywhere from five to 15 contacts per day. And, if you did that, you would then have to run a Monte Carlo simulation on the model that would produce distributions of values. For the purpose of this explanation, we'll just stick with fixed values, because the other advantage of Analytica is that it allows you to do a little bit of parametric analysis at the end. In some cases, it allows you a contact rate reduction. If you click on that, you'll see a whole range of values.

In some ways, when you select all, you're telling the model that you want to test out every single combination of contact rate and produce a result that the user will test out. For example, what implication does it have if contact rate was 20% less versus 40%, etcetera? You'll have access to all those scenarios when you view the results. In some ways, you are capturing a combination of both uncertainty and a little bit of a scenario analysis. To test - or just to explore that a little bit, we'll look at some of the results. First one would be, for example, the daily incidents.

If you click on that, you get your very traditional epidemic curve that most of you are probably reasonably familiar with. It shows the daily incidents of cases as time progresses with a certain number of input or a certain combination of input parameters. We've assumed that they were looking at a 20% reduction in contact rate, and, as I mentioned a few seconds ago, you can look at alternative contact rate reduction. In this particular case, we've got - if we put zero here - the incidents if contacts were maintained at their standard rate level. Zero. You can look at a 10% , 20% or 30% reduction. This shows the effect of changing contact rates in the general community and what implications that would have on the daily incidents of disease.

What's perhaps more interesting or easier to understand, you could potentially just plot them all on one curve. You've got your original incidents curve and the implications associated with changing contact rates. That allows you to have a quick look at what effect contact rate reduction has

on the overall incidents curve. Here, you can clearly see a reduction in their peak and a spreading out. You're reducing the highest number of cases, or the highest number of incidents that could occur in any one day, essentially stretching it out over a longer period of time.

In addition to contact rate reduction, you could look at other things. For example, what would be the implications of changing the duration of time it takes for a test result? How much of an effect does that have on the incidents curve? So, if I change the plot over here, you can see I changed it to time for test results. You might say, "Well, that didn't really do anything," and the reason it hasn't done anything is because the proportion of cases detected is currently set at zero. Let's change that to 10% or 20% to see what effect that now has. It shows you the effect on the epidemic curve if you were detecting 20% of your cases and you were able to get test results within one day, two days, three days, four days, etcetera.

The interpretation there, obviously, is that if you are able to capture infected individuals and isolate them in a rapid time. They are no longer in a position to spread infection within the community in general. Therefore, you start to bring that curve down. All of that is assuming you have 20% of the portion of cases detected with no contact tracing going on. Each curves shows the time for each test result. That's what would happen with no contact tracing. If you implement contact tracing, it adds additional effect. For example, as you add more and more contact tracing, it only improves the process further. It's a relatively simple initial pass model.

There will also be a relationship between your ability to contact trace and your ability to test. What the model currently doesn't have, or at least this model doesn't have, is the relationship between the two. They're handled independently and it's up to the user to make sure that they're entering a logical value. On that note, there's nothing in the model that would tell me that I can't have 100% contact tracing and 0% testing. But, logically, that would make absolutely no sense. In essence, you know, I could have a setting in here that says, "zero proportion of the cases are detected," and set time to contact trace. The time for test result here is set to seven, contract trace effect on this is that 10% proportion of cases detected is zero.

This setting is telling you that I am detecting no cases. So, my testing is virtually useless. My contact trace effectiveness is at 10%. Or, let's go to 20% just to make it more dramatic. Contact tracing is at 20% and time to contact trace are each of these curves. Now, logic would tell you that if you are detecting no cases, it's highly unlikely - well, virtually impossible - that you'd be contact tracing 30% of the population. The model isn't meant to link testing a certain percentage in order to effectively trace a certain other proportion. That's left up to the user of this relatively simple model to interpret the results with caution. That's what I want to convey in this portion - you need to have use some logic when you test out various scenarios.

This would be a prime example of the classic garbage in, garbage out. You can come up with combinations of values that don't make sense from a real world perspective. You should always interpret the results coming out with a little bit of common sense and public-health insight. Having said that, there are various other results that you can generate. We just went through daily incidents. You can produce results that show the total attack rate. That's essentially a cumulative curve that shows the proportion of the total population effected at the end of the pandemic. So, similar to the previous curves but just showing the final attack rate. This, too, you can plot using different curves.

Let's say we wanted to see what the testing effect had on the final attack rate. You could look at time test results. The contact rate reduction here, we can set that to 20%. We'll eliminate contact tracing, so contact trace is set to zero. Time to contact trace is seven. Proportion of cases detected, let's just reduce that a little bit to 30%. So, we have a testing program which has various time delays that we want to explore, and we want to see what impact that has on the final attack rate. You can see here with the pink line, that's seven days, and after 365 days, you would end up with almost 65% of the Canadian population effected. This would be you doing nothing, but the contact rate is reduced by 20%. So, people are socially distancing a little bit. You are doing nothing else except, in 30% of the time, you are isolating sick people.

Now, this is probably not a realistic scenario either. But, it helps to understand, if my only tool was speeding up testing time, and I kept it at 30%, I'm still only capturing 30%. Those 30% get isolated so, after a certain number of days, I find

them and I stop them from spreading infection more. That's what the effect would be on the final attack rate.

I could have a pretty significant reduction going from more than 50% down to under 10% just by speeding up the process. Obviously, the next stage would be to consider if it's feasibly possible to do testing and isolation in a one day turnaround? That becomes the next challenge. At least you get some insight as to whether this is something worth exploring further, and then we can figure out how to actually make it happen.

So, that's attack rate. And, similarly, you've got various other inputs and outputs that you can explore - that's all of these buttons. I will point out, in fact, that if you are to open up the actual model, and you wanted to look at the result for any of these bubbles, you can also do that. For example, if I clicked on the exposed, you'll see below here, I get two little additional bubbles. If I click on the green one, that's a result. That shows me the result for that particular parameter looks like for various settings. So, the results in your front-end screen here, these results, are not restricted to only producing those results.

If you are interested in looking at any other intermediate parameter, all you have to do is open up the model, click on the value once, and you can click on the green button down here or up on the top is a similar "show result." If you click on that, it'll show you the mid-value for what symptomatic severe looks like. So, whatever that parameter is, it'll show you what that looks like. You can similarly play around with changing every other parameter that has an influence on this parameter. It will show up on the top of your screen and you'll have the ability to modify it and see what result and impact it has on the result.

The underlying message, as I said earlier on, though, is you should always put some caution and critical though into the results. The model doesn't restrict you. You can produce some combinations that don't make practical sense. Always have that in the back of your mind. That would conclude the presentation of the PHAC compartmental model constructed using Analytica software. Thanks.

Margaret Harris Brockman: Thanks very much Aamir for this presentation and demonstrating how to use the model. We're now going to begin our live question and answer period with Aamir. So,

Aamir, our first question that came up in the course of your presentation was, "What were your data sources in developing the original model? Have you been rerunning the model as the epidemic dynamics change in Canada?"

Aamir Fazil: Thanks. At the onset, a lot of the information was derived from the published literature. We have a team at PHAC that was combing through both pre-publication literature to derive as best as we can some of the input parameters. As I said, that has constantly evolved as time has gone on. This model was intended to serve as the initial template, if you will, to think about what needs to be captured and how the trajectory might look.

We developed this model, used initial information from the published literature, and what was known at the time in Canada. Eventually, this model has started to be used less and we've moved into deploying other models that are a lot more suitable to purpose, and they'll be shared on the NCCID site eventually. That includes agent based modelling approaches, and a more complex SEIR model with age stratification that has been fit to epidemic, real-world surveillance data.

Margaret Harris Brockman: OK, thank you. The next question that arises is, "are provinces or territories using their own models and are they similar or different from the model and the impact on the modelled outcomes that you've been talking about?" One example, I guess, is Saskatchewan's compartmental model.

Aamir Fazil: I think most provinces have their own models that they've developed internally. We also have been, across Canada, partnering and sharing models and ideas. So, in certain situations like the mentioned agent based model, we've found it applicable in other provinces as the needs have arisen. For the most part, many provinces have developed their own models. When requested, models have been shared to help in the decision-making process.

As we all recognize, the parameters and the nature of the epidemic in different provinces requires a little bit more nuance in terms of the population, which would be very different in each province. Taking a very federal view, you can't necessarily go down to that level of disaggregation. It's very important, I think, that most of the modelling or the insights are gained at a provincial and territorial level, if not even lower level than that.

Margaret Harris Brockman: Thanks, Aamir. I think that leads into the next question, "if I manipulate the input parameters, or other features, as you've been describing in Analytica, will I be able to save them for our use, and what about changes to the model itself?"

Aamir Fazil: Sorry, could you repeat that again, Margaret?

Margaret Harris Brockman: "If I manipulate the input parameters, or other features, as you were describing, in Analytica, will I be able to save it for our own use and what about changes made to the model itself?"

Aamir Fazil: Unfortunately, Analytica is a propriety software. So, with the free viewer, you're able to look at the model, run the model, and save the results. You could change it, put parameters, and produce results that correspond to those input parameters. However, with the free version, I believe you are restricted in terms of not being able to change the equations or the structure of the model.

That would be the only restriction. For the most part, the intent here is to be able to view it, for free.

I believe everyone can see my screen at this time. When you open up the details of the model, there is - if there was an absolute desire to recreate the model in, for example, another software you have available, then there should be enough information in this screen, including all the necessary equations, that would allow you to recreate it if you had to.

However, this was a model designed to give you initial insights into the potential trajectory and the potential impact of various interventions. So, I can't imagine there would be much need to dramatically modify the model. There might be a need to test out alternative scenarios and look at what that looks like. Then, you would probably be encouraged to develop a model that is a lot more detailed and perhaps more specific to the region or the location to which the model is being applied. It would also likely include considerations like age stratification. We will share some of those models soon. Like I said, another SEIR model, which has age stratification.

Those might be more appropriate for that purpose. I think this model is most appropriate to get a quick overview or a quick insight into what the impact of various interventions

are, could be, and how significant the trajectory might look in your particular area, depending upon various assumptions.

Margaret Harris Brockman: Thank you, Aamir. Is there anything else you would like to add? This particular presentation was recorded a little bit earlier and I know that there's often a second thought about anything you think would be of interest to the audience.

Aamir Fazil: Yeah, I think I touched on the idea of entering uncertainty distributions. I think that's an important consideration too. Whenever we produce results from models, at least the ones that I demonstrated here, a curve looks like this. The epidemic curve is simply a best estimate based on the individual best intimate input parameters. Ultimately, all those parameters have uncertainty and, in honesty, variability associated with them. It would be best practice, as I mentioned, to have probability distributions defined for your input parameters.

Once you have those probability distributions defined, you would have to go under uncertainty options here. Under sample size, you would define how many iterations you want the model to run. If you entered 1,000, for example, the model would sample from each distribution, run the model, produce results, and resample from the probability distributions. Ultimately, you'd end up with a result that would not be just a simple line. You'd, hopefully, end up with a cloud of epidemic curves that you can interpret to represent the range of results that you would expect, given the uncertainty and the input distributions. Now, it's important to recognize that that is, again, really just an uncertainty distribution.

It's not the stochastic model that many of you might be thinking. When you run an agent-based model, there are very stochastic elements associated with them that represent the epidemic dying out or probability events that wouldn't necessarily be captured in a model like this. But, at least by putting probability distributions, it allows you to recognize that there is a range of possibilities driven by our uncertainty in those input parameters. As we've been discussing, the nature of information in this particular outbreak is evolving daily and weekly. It's important to capture that uncertainty.

Margaret Harris Brockman: Yes, evolving almost hourly it seems sometimes. That has brought us to the end of the question and answer period. I would like to take a moment to thank you, Aamir Fazil, for

joining us again. It was an excellent seminar, and we really appreciate the time you've taken today, not only for this presentation and webinar, but also working with NCCID to develop the additional materials. You can find information related to mathematical modelling from Public Health and this SEIR model on our website, NCCID.ca.

This webinar is part of a series that NCCID is sharing on mathematical modelling and COVID-19, the *Synergy Series*. The next webinar will be September 11th and is titled "COVID-19 Vaccine: Outlook and Opportunities for Modelling with Dr. Joanne Langly." For information on all of these and to obtain a copy of the recording of this presentation, you can visit us at NCCID.ca or @CentreInfection on Twitter. Please stay well and safe. Thank you.